```
Qy       661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2197 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2256

Qy       681 ProLeuLysGluTyrPheAlaCysVal 689
            |||||||||||||||||||||||||||
Db      2257 CCGCTGAAGGAGTATTTTGCGTGTGTG 2283

RESULT 2
S-10-144-577-2
 Sequence 2, Application US/10144577
 Publication No. US20030083292A1
 GENERAL INFORMATION:
  APPLICANT: MacLeod, Alan Robert
  TITLE OF INVENTION: Inhibitors of DNA Methyltransferase Isoforms
  FILE REFERENCE: MET-005
  CURRENT APPLICATION NUMBER: US/10/144,577
  CURRENT FILING DATE:  2002-05-13
  PRIOR APPLICATION NUMBER: US 60/290,202
  PRIOR FILING DATE: 2001-05-11
  PRIOR APPLICATION NUMBER: US 60/290,212
  PRIOR FILING DATE: 2001-05-11
  NUMBER OF SEQ ID NOS: 49
  SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 2
   LENGTH: 3005
   TYPE: DNA
   ORGANISM: Homo sapiens
S-10-144-577-2

Alignment Scores:
Seq. No.:                0          Length:          3005
Score:                   3763.00    Matches:         689
Percent Similarity:      100.0%     Conservative:    0
Best Local Similarity:   100.0%     Mismatches:      0
Query Match:             100.0%     Indels:          0
DB:                      6          Gaps:            0

US-10-623-813-86 (1-689) x US-10-144-577-2 (1-3005)

Qy         1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       899 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 958

Qy        21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       959 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1018

Qy        41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1019 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1078

Qy        61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1079 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1138

Qy        81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1139 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1198

Qy       101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1199 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1258

Qy       121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1259 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1318

Qy       141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1319 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1378

Qy       161 ProProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1379 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1438

Qy       181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1439 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1498

Qy       201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1499 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1558

Qy       221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1559 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1618

Qy       241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1619 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1678

Qy       261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1679 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1738

Qy       281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1739 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1798

Qy       301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1799 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1858

Qy       321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1859 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1918

Qy       341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1919 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1978

Qy       361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1979 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2038

Qy       381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2039 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2098

Qy       401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2099 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2158

Qy       421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2159 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2218

Qy       441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2219 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2278

Qy       461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2279 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2338

Qy       481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2339 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2398

Qy       501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2399 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2458

Qy       521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
```

```
Db       2459 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2518

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560

Db       2519 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2578

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580

Db       2579 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2638

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600

Db       2639 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2698

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620

Db       2699 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2758

Qy        621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640

Db       2759 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2818

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660

Db       2819 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2878

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680

Db       2879 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2938

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689

Db       2939 CCGCTGAAGGAGTATTTTGCGTGTGTG 2965
```

RESULT 3
US-10-172-118-132
Sequence 132, Application US/10172118
[illegible]

```
Db      2039 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2098

Qy       401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2099 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2158

Qy       421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2159 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2218

Qy       441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2219 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2278

Qy       461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2279 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2338

Qy       481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2339 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2398

Qy       501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2399 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2458

Qy       521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2459 AAGGAGGGAGATGATCGCCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2518

Qy       541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2519 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2578

Qy       561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2579 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2638

Qy       581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2639 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2698

Qy       601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2699 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2758

Qy       621 LysAspGlnHisProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2759 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2818

Qy       641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2819 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2878

Qy       661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2879 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2938

Qy       681 ProLeuLysGluTyrPheAlaCysVal 689
             |||||||||||||||||||||||||||
Db      2939 CCGCTGAAGGAGTATTTTGCGTGTGTG 2965

RESULT 3
AF067972
LOCUS       AF067972                3005 bp    mRNA    linear   PRI 12-FEB-2001
DEFINITION  Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA,
            complete cds.
ACCESSION   AF067972
            AF067972.2  GI:12746531

            sapiens (human)
            ens
            Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 3005)
  AUTHORS   Xie,S., Wang,Z., Okano,M., Nogami,M., Li,Y., He,W.W., Okumura,K.
            and Li,E.
  TITLE     Cloning, expression and chromosome locations of the human DNMT3
            gene family
  JOURNAL   Gene 236 (1), 87-95 (1999)
  PUBMED    10433969
REFERENCE   2  (bases 1 to 3005)
  AUTHORS   Xie,S., Okano,M. and Li,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-MAY-1998) CVRC, Mass. General Hospital, 149 13th
            Street, Charlestown, MA 02129, USA
REFERENCE   3  (bases 1 to 3005)
  AUTHORS   Hata,K., Shirohzu,H., Sasaki,H. and En,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-FEB-2001) CVRC, Mass. General Hospital, 149 13th
            Street, Charlestown, MA 02129, USA
  REMARK    Sequence update by submitter
COMMENT     On Feb 12, 2001 this sequence version replaced gi:4927369.
FEATURES             Location/Qualifiers
     source          1. .3005
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2p23; between D2S171 and D2S174; 48-50 cM"
     gene            1. .3005
                     /gene="DNMT3A"
     CDS             230. .2968
                     /gene="DNMT3A"
                     /function="de novo DNA cytosine methyltransferase"
                     /codon_start=1
                     /product="DNA cytosine methyltransferase 3 alpha".
                     /protein_id="AAD33084.2"
                     /db_xref="GI:12746532"
                     /translation="MPAMPSSGPGDTSSSAAEREEDRKDGEEQEEPRGKEERQEPSTT
                     ARKVGRPGRKRKHPPVESGDTPKDPAVISKSPSMAQDSGASELLPNGDLEKRSEPQPE
                     EGSPAGGQKGGAPAEGEGAAETLPEASRAVENGCCTPKEGRGAPAEAGKEQKETNIES
                     MKMEGSGRGRLRGGLGWESSLRQRPMPRLTFQAGDPYYISKRKRDEWLARWKREAEKKA
                     KVIAGMNAVEENQGPGESQKVEEASPPAVQQPTDPASPTVATTPEPVGSDAGDKNATK
                     AGDDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDG
                     KFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPVCHDSDE
                     SDTAKAVEVQNKPMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYAP
                     PPPAKKPRKSTAEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPL
                     FVGGMCQNCKNCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLL
                     VGPGAAQAAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYP
                     PVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIM
                     YVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHD
                     ARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNL
                     PGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKE
                     DILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV"
ORIGIN

Alignment Scores:
Pred. No.:                0                Length:          3005
Score:                    3763.00          Matches:         689
Percent Similarity:       100.0%           Conservative:    0
Best Local Similarity:    100.0%           Mismatches:      0
Query Match:              100.0%           Indels:          0
DB:                       5                Gaps:            0

US-10-623-813-86 (1-689) x AF067972 (1-3005)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 958

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1018
```

```
y      41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
b    1019 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1078
y      61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
b    1079 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1138
y      81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
b    1139 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1198
y     101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
b    1199 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1258
y     121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
b    1259 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1318
y     141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
b    1319 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1378
y     161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
b    1379 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1438
y     181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
b    1439 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1498
y     201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
b    1499 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1558
y     221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
b    1559 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1618
y     241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
b    1619 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1678
y     261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
b    1679 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1738
y     281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
b    1739 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1798
y     301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
b    1799 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1858
y     321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db   1859 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1918
Qy    341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
Db   1919 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1978
Qy    361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
Db   1979 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2038
Qy    381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db   2039 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2098
Qy    401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420

Db   2099 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2158
Qy    421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db   2159 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2218
Qy    441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db   2219 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2278
Qy    461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db   2279 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2338
Qy    481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db   2339 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2398
Qy    501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db   2399 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2458
Qy    521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
Db   2459 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2518
Qy    541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db   2519 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2578
Qy    561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db   2579 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2638
Qy    581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
Db   2639 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2698
Qy    601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db   2699 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2758
Qy    621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db   2759 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2818
Qy    641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db   2819 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2878
Qy    661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db   2879 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2938
Qy    681 ProLeuLysGluTyrPheAlaCysVal 689
Db   2939 CCGCTGAAGGAGTATTTTGCGTGTGTG 2965
```

RESULT 4
CS050242
LOCUS CS050242 3604 bp DNA linear PAT 23-MAR-2005
DEFINITION Sequence 26 from Patent WO2005021757.
ACCESSION CS050242
VERSION CS050242.1 GI:61889497
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Tomme,P.H. and van Rompaey,L.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 5, 2006, 18:12:34 ; Search time 10519 Seconds
(without alignments)
6282.884 Million cell updates/sec

Title: US-10-623-813-86
Perfect score: 3763
Sequence: 1 MNAVEENQGPGESQKVEEAS.....WSVPVIRHLFAPLKEYFACV 689

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10623813/runat_05102006_121146_27659/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05h
-USER=US10623813_@CGN_1_1_6706_@runat_05102006_121146_27659 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_sts:*
8: gb_sy:*
9: gb_un:*
10: gb_vi:*
11: gb_ov:*
12: gb_htg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3763	100.0	2371	5	AF480163	AF480163 Homo sapi
2	3763	100.0	3005	2	CS050243	CS050243 Sequence
3	3763	100.0	3005	5	AF067972	AF067972 Homo sapi
4	3763	100.0	3604	2	CS050242	CS050242 Sequence
5	3763	100.0	4258	5	AF331856	AF331856 Homo sapi
6	3763	100.0	4294	2	DD171352	DD171352 Marker fo
7	3763	100.0	4294	5	BC043617	BC043617 Homo sapi
8	3763	100.0	4395	2	CS050241	CS050241 Sequence
9	3746	99.5	2841	2	CQ720469	CQ720469 Sequence
10	3714	98.7	2318	6	AF480164	AF480164 Mus muscu
11	3714	98.7	4094	6	BC007466	BC007466 Mus muscu
12	3714	98.7	4192	6	AF068625	AF068625 Mus muscu
13	3712	98.6	2597	14	AY271299	AY271299 Bos tauru
14	3712	98.6	2798	14	AY271298	AY271298 Bos tauru
15	3702	98.4	4476	5	AB208833	AB208833 Homo sapi
16	3570.5	94.9	3014	11	AB214886	AB214886 Gallus ga
c 17	3226	85.7	264380	12	AC120824	AC120824 Rattus no
c 18	3085.5	82.0	253474	12	AC095441	AC095441 Rattus no
19	3076	81.7	3377	11	AB196917	AB196917 Danio rer
20	3024.5	80.4	235547	12	AC112586	AC112586 Rattus no
21	2944	78.2	2933	11	AB196919	AB196919 Danio rer
c 22	2447.5	65.0	344787	12	AC112040	AC112040 Rattus no
23	2413	64.1	2191	5	AK025230	AK025230 Homo sapi
24	2292.5	60.9	3544	11	AB196915	AB196915 Danio rer
25	2204.5	58.6	168651	5	AC009474	AC009474 Homo sapi
26	2174.5	57.8	4135	6	AF068627	AF068627 Mus muscu
27	2172.5	57.7	4163	6	AF151974	AF151974 Mus muscu
28	2172.5	57.7	4278	6	AF151970	AF151970 Mus muscu
29	2169.5	57.7	4195	6	AF068626	AF068626 Mus muscu
30	2167.5	57.6	4223	6	AF151973	AF151973 Mus muscu
31	2167.5	57.6	4338	6	AF151969	AF151969 Mus muscu
32	2165	57.5	4267	5	AF176228	AF176228 Homo sapi
33	2159	57.4	3630	6	BC105922	BC105922 Mus muscu
34	2157	57.3	4043	2	CQ726303	CQ726303 Sequence
35	2157	57.3	4145	5	AF156488	AF156488 Homo sapi
36	2157	57.3	4335	2	CQ981515	CQ981515 Sequence
37	2157	57.3	4335	5	AF331857	AF331857 Homo sapi
38	2149.5	57.1	4369	5	DQ321787	DQ321787 Homo sapi
39	2062	54.8	2491	14	AY244710	AY244710 Bos tauru
40	2012.5	53.5	6248	11	AB196918	AB196918 Danio rer
41	1948.5	51.8	2421	14	AY244712	AY244712 Bos tauru
42	1902	50.5	3946	6	AF068628	AF068628 Mus muscu
43	1900	50.5	3974	6	AF151976	AF151976 Mus muscu
44	1900	50.5	4089	6	AF151971	AF151971 Mus muscu
45	1897	50.4	4006	6	AY078427	AY078427 Mus muscu

ALIGNMENTS

RESULT 1
AF480163
LOCUS AF480163 2371 bp mRNA linear PRI 14-OCT-2002
DEFINITION Homo sapiens DNA cytosine methyltransferase 3A2 (DNMT3A2) mRNA, complete cds.
ACCESSION AF480163
VERSION AF480163.1 GI:23954437
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2371)
AUTHORS Chen,T., Ueda,Y., Xie,S. and Li,E.
TITLE A Novel Dnmt3a Isoform Produced from an Alternative Promoter Localizes to Euchromatin and Its Expression Correlates with Active de Novo Methylation
JOURNAL J. Biol. Chem. 277 (41), 38746-38754 (2002)
PUBMED 12138111
REFERENCE 2 (bases 1 to 2371)
AUTHORS Chen,T. and Li,E.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2002) Cardiovascular Research Center, Massachusetts General Hospital, 149 13th Street, Charlestown, MA 02129, USA
FEATURES Location/Qualifiers

```
     source          1. .2371
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2p23"
     gene            1. .2371
                     /gene="DNMT3A2"
     CDS             217. .2286
                     /gene="DNMT3A2"
                     /function="de novo DNA cytosine methylation"
                     /codon_start=1
                     /product="DNA cytosine methyltransferase 3A2"
                     /protein_id="AAN40037.1"
                     /db_xref="GI:23954438"
                     /translation="MNAVEENQGPGESQKVEEASPPAVQQPTDPASPTVATTPEPVGS
                     DAGDKNATKAGDDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGT
                     RWVMWFGDGKFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKL
                     FPVCHDSDESDTAKAVEVQNKPMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWV
                     EPEAAAYAPPPPAKKPRKSTAEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGS
                     LNVTLEHPLFVGGMCQNCKNCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRC
                     FCVECVDLLVGPGAAQAAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQ
                     EFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVG
                     MVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLF
                     FEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAH
                     RARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQH
                     FPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPL
                     KEYFACV"
ORIGIN

Alignment Scores:
Pred. No.:              0               Length:         2371
Score:                  3763.00         Matches:        689
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     5               Gaps:           0

US-10-623-813-86 (1-689) x AF480163 (1-2371)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 276

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 336

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 396

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 456

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 516

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 576

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 636

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 696

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 756

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 816

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 876

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 936

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTC 996

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1056

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1057 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1116

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1176

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1177 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1236

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1237 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1296

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 1356

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1357 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 1416

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1417 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 1476

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 1536

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1537 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 1596

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 1656

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1657 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 1716

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1717 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 1776

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1777 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 1836

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1837 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 1896

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1897 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 1956

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1957 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2016

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2017 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2076

Qy        621 LysAspGlnHisProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2077 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2136

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2137 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2196

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2197 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2256

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              ||||||||||||||||||||||||||
Db       2257 CCGCTGAAGGAGTATTTTGCGTGTGTG 2283

RESULT 2
CS050243
LOCUS       CS050243                3005 bp    DNA     linear   PAT 23-MAR-2005
DEFINITION  Sequence 27 from Patent WO2005021757.
ACCESSION   CS050243
VERSION     CS050243.1  GI:61889498
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1.
  AUTHORS   Tomme,P.H. and van Rompaey,L.
  TITLE     Polypeptides and polynucleotides for use as a medicament
  JOURNAL   Patent: WO 2005021757-A 27 10-MAR-2005;
            Galapagos Genomics N.V. (BE)
FEATURES             Location/Qualifiers
     source          1. .3005
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          3005
Score:                    3763.00        Matches:         689
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       2              Gaps:            0

US-10-623-813-86 (1-689) x CS050243 (1-3005)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 958

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1018

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1078

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1138

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1198

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1258

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1318

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1378

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1438

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1439 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1498

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1558

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1559 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1618

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1619 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1678

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1679 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1738

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1739 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1798

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1799 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1858

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1859 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1918

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1919 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1978

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1979 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2038

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2039 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2098

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2099 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2158

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2159 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2218

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2278

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2279 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2338

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2339 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2398

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2399 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2458

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2459 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2518

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2519 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2578

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2579 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2638

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2639 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2698

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2699 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2758

Qy        621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2759 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2818

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2819 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2878

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2879 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2938

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              ||||||||||||||||||||||||||||
Db       2939 CCGCTGAAGGAGTATTTTGCGTGTGTG 2965

RESULT 3
AF067972
LOCUS       AF067972               3005 bp    mRNA    linear   PRI 12-FEB-2001
DEFINITION  Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA,
            complete cds.
ACCESSION   AF067972
VERSION     AF067972.2  GI:12746531
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 3005)
  AUTHORS   Xie,S., Wang,Z., Okano,M., Nogami,M., Li,Y., He,W.W., Okumura,K.
            and Li,E.
  TITLE     Cloning, expression and chromosome locations of the human DNMT3
            gene family
  JOURNAL   Gene 236 (1), 87-95 (1999)
   PUBMED   10433969
REFERENCE   2  (bases 1 to 3005)
  AUTHORS   Xie,S., Okano,M. and Li,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-MAY-1998) CVRC, Mass. General Hospital, 149 13th
            Street, Charlestown, MA 02129, USA
REFERENCE   3  (bases 1 to 3005)
  AUTHORS   Hata,K., Shirohzu,H., Sasaki,H. and En,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-FEB-2001) CVRC, Mass. General Hospital, 149 13th
            Street, Charlestown, MA 02129, USA
  REMARK    Sequence update by submitter
COMMENT     On Feb 12, 2001 this sequence version replaced gi:4927369.
FEATURES             Location/Qualifiers
     source          1. .3005
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2p23; between D2S171 and D2S174; 48-50 cM"
     gene            1. .3005
                     /gene="DNMT3A"
     CDS             230. .2968
                     /gene="DNMT3A"
                     /function="de novo DNA cytosine methyltransferase"
                     /codon_start=1
                     /product="DNA cytosine methyltransferase 3 alpha"
                     /protein_id="AAD33084.2"
                     /db_xref="GI:12746532"
                     /translation="MPAMPSSGPGDTSSSAAEREEDRKDGEEQEEPRGKEERQEPSTT
                     ARKVGRPGRKRKHPPVESGDTPKDPAVISKSPSMAQDSGASELLPNGDLEKRSEPQPE
                     EGSPAGGQKGGAPAEGEGAAETLPEASRAVENGCCTPKEGRGAPAEAGKEQKETNIES
                     MKMEGSRGRLRGGLGWESSLRQRPMPRLTFQAGDPYYISKRKRDEWLARWKREAEKKA
                     KVIAGMNAVEENQGPGESQKVEEASPPAVQQPTDPASPTVATTPEPVGSDAGDKNATK
                     AGDDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDG
                     KFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPVCHDSDE
                     SDTAKAVEVQNKPMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYAP
                     PPPAKKPRKSTAEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPL
                     FVGGMCQNCKNCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLL
                     VGPGAAQAAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYP
                     PVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIM
                     YVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHD
                     ARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNL
                     PGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKE
                     DILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV"
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:          3005
Score:                  3763.00        Matches:         689
Percent Similarity:     100.0%         Conservative:    0
Best Local Similarity:  100.0%         Mismatches:      0
Query Match:            100.0%         Indels:          0
DB:                     5              Gaps:            0

US-10-623-813-86 (1-689) x AF067972 (1-3005)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 958

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1018
```

```
Qy        41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
Db      1019 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1078

Qy        61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
Db      1079 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1138

Qy        81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
Db      1139 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1198

Qy       101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
Db      1199 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1258

Qy       121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db      1259 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1318

Qy       141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db      1319 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1378

Qy       161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
Db      1379 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1438

Qy       181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db      1439 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1498

Qy       201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db      1499 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1558

Qy       221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
Db      1559 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1618

Qy       241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db      1619 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1678

Qy       261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db      1679 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1738

Qy       281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db      1739 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1798

Qy       301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1799 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1858

Qy       321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db      1859 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1918

Qy       341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
Db      1919 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1978

Qy       361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
Db      1979 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2038

Qy       381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db      2039 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2098

Qy       401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      2099 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2158

Qy       421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      2159 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2218

Qy       441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      2219 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2278

Qy       461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db      2279 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2338

Qy       481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db      2339 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2398

Qy       501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      2399 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2458

Qy       521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
Db      2459 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2518

Qy       541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db      2519 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2578

Qy       561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db      2579 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2638

Qy       581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
Db      2639 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2698

Qy       601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db      2699 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2758

Qy       621 LysAspGlnHisProProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db      2759 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2818

Qy       641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2819 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2878

Qy       661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db      2879 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2938

Qy       681 ProLeuLysGluTyrPheAlaCysVal 689
Db      2939 CCGCTGAAGGAGTATTTTGCGTGTGTG 2965
```

```
RESULT 4
CS050242
LOCUS       CS050242               3604 bp    DNA     linear   PAT 23-MAR-2005
DEFINITION  Sequence 26 from Patent WO2005021757.
ACCESSION   CS050242
VERSION     CS050242.1  GI:61889497
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tomme,P.H. and van Rompaey,L.
```

```
  TITLE     Polypeptides and polynucleotides for use as a medicament
  JOURNAL   Patent: WO 2005021757-A 26 10-MAR-2005;
            Galapagos Genomics N.V. (BE)
FEATURES             Location/Qualifiers
     source          1. .3604
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              0               Length:         3604
Score:                  3763.00         Matches:        689
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     2               Gaps:           0

US-10-623-813-86 (1-689) x CS050242 (1-3604)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 276

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 336

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 396

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 456

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 516

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 576

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 636

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 696

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 756

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 816

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 876

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 936

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 996

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1056

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1057 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1116

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1176

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1177 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1236

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1237 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1296

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 1356

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1357 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 1416

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1417 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 1476

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 1536

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1537 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 1596

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 1656

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1657 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 1716

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1717 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 1776

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1777 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 1836

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1837 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 1896

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1897 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 1956

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1957 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2016

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2017 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2076

Qy        621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2077 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2136

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2137 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2196

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2197 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2256

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              |||||||||||||||||||||||||||
Db       2257 CCGCTGAAGGAGTATTTTGCGTGTGTG 2283

RESULT 5
AF331856
LOCUS       AF331856                4258 bp    mRNA    linear   PRI 02-JAN-2002
DEFINITION  Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA,
            complete cds.
ACCESSION   AF331856
VERSION     AF331856.1  GI:18033252
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 4258)
  AUTHORS   Ni,J., Pradhan,S. and Roberts,R.J.
  TITLE     Cloning, expression and characterization of human DNMT3 genes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4258)
  AUTHORS   Ni,J., Pradhan,S. and Roberts,R.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-DEC-2000) New England Biolabs, 32 Tozer Road,
            Beverly, MA 01915, USA
FEATURES             Location/Qualifiers
     source          1. .4258
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
     gene            1. .4258
                     /gene="DNMT3A"
     CDS             227. .2956
                     /gene="DNMT3A"
                     /codon_start=1
                     /product="DNA cytosine methyltransferase 3 alpha"
                     /protein_id="AAL57039.1"
                     /db_xref="GI:18033253"
                     /translation="MPSSGPGDTSSSAAEREEDRKDGEEQEEPRGKEERQEPSTTARK
                     VGRPGRKRKHPPVESGDTPKDPAVISKSPSMAQDSGASELLPNGDLEKRSEPQPEEGS
                     PAGGQKGGAPAEGEGAAETLPEASRAVENGCCTPKEGRGAPAEAGKEQKETNIESMKM
                     EGSRGRLRGGLGWESSLRQRPMPRLTFQAGDPYYISKRKRDEWLARWKREAEKKAKVI
                     AGMNAVEENQGPGESQKVEEASPPAVQQPTDPASPTVATTPEPVGSDAGDKNATKAGD
                     DEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFS
                     VVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPVCHDSDESDT
                     AKAVEVQNKPMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYAPPPP
                     AKKPRKSTAEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPLFVG
                     GMCQNCKNCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLLVGP
                     GAAQAAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYPPVP
                     AEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVG
                     DVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARP
                     KEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGM
                     NRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDIL
                     WCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV"
ORIGIN

Alignment Scores:
Pred. No.:              0           Length:        4258
Score:                  3763.00     Matches:       689
Percent Similarity:     100.0%      Conservative:  0
Best Local Similarity:  100.0%      Mismatches:    0

Query Match:            100.0%      Indels:        0
DB:                     5           Gaps:          0

US-10-623-813-86 (1-689) x AF331856 (1-4258)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 946

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1006

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1066

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1126

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1186

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1246

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1306

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1366

Qy        161 ProProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1426

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1486

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1546

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1606

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1607 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1666

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1667 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1726

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1727 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1786

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1787 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1846

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1847 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1906
```

```
Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1907 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1966

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1967 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2026

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2027 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2086

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2087 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2146

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2147 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2206

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2207 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2266

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2267 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2326

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2327 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2386

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2387 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2446

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2447 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2506

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2507 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2566

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2567 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2626

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2627 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2686

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2687 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2746

Qy        621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2747 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2806

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2807 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2866

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2867 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2926

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              ||||||||||||||||||||||||||||
Db       2927 CCGCTGAAGGAGTATTTTGCGTGTGTG 2953

RESULT 6
```

```
DD171352
LOCUS       DD171352                4294 bp    DNA     linear   PAT 19-DEC-2005
DEFINITION  Marker for diagnosing human subarachnoid hemorrhage and its use.
ACCESSION   DD171352
VERSION     DD171352.1  GI:83959786
KEYWORDS    JP 2005151854-A/10.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 4294)
  AUTHORS   Yamaura,A., Kubota,M., Machida,T., Uchino,F., Kobayashi,E.,
            Saeki,N., Takiguchi,M. and Hiwasa,T.
  TITLE     Marker for diagnosing human subarachnoid hemorrhage and its use
  JOURNAL   Patent: JP 2005151854-A 10 16-JUN-2005;
            Japan Science and Technology Agency
COMMENT     OS   Homo sapiens
            PN   JP 2005151854-A/10
            PD   16-JUN-2005
            PF   21-NOV-2003 JP 2003393161
            PI   akira yamaura,motoo kubota,toshio machida,fukuo uchino,eiichi
            PI   kobayashi,
            PI   naokatsu saeki,masaki takiguchi,takaki hiwasa CC
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .4294
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:          4294
Score:                  3763.00        Matches:         689
Percent Similarity:     100.0%         Conservative:    0
Best Local Similarity:  100.0%         Mismatches:      0
Query Match:            100.0%         Indels:          0
DB:                     2              Gaps:            0

US-10-623-813-86 (1-689) x DD171352 (1-4294)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 966

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1026

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1086

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1146

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1206

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1266

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1326

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1386
```

```
Qy       161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
Db      1387 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1446

Qy       181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db      1447 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1506

Qy       201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db      1507 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1566

Qy       221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
Db      1567 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1626

Qy       241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db      1627 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1686

Qy       261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db      1687 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1746

Qy       281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db      1747 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1806

Qy       301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1807 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1866

Qy       321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db      1867 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1926

Qy       341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
Db      1927 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1986

Qy       361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
Db      1987 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2046

Qy       381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db      2047 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2106

Qy       401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      2107 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2166

Qy       421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      2167 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2226

Qy       441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      2227 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2286

Qy       461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db      2287 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2346

Qy       481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db      2347 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2406

Qy       501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      2407 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2466

Qy       521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
Db      2467 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2526

Qy       541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db      2527 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2586

Qy       561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db      2587 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2646

Qy       581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
Db      2647 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2706

Qy       601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db      2707 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2766

Qy       621 LysAspGlnHisProProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db      2767 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2826

Qy       641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2827 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2886

Qy       661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db      2887 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2946

Qy       681 ProLeuLysGluTyrPheAlaCysVal 689
Db      2947 CCGCTGAAGGAGTATTTTGCGTGTGTG 2973
```

RESULT 7
BC043617

```
LOCUS       BC043617               4294 bp    mRNA    linear   PRI 30-JUN-2004
DEFINITION  Homo sapiens DNA (cytosine-5-)-methyltransferase 3 alpha,
            transcript variant 3, mRNA (cDNA clone MGC:50948 IMAGE:6150112),
            complete cds.
ACCESSION   BC043617
VERSION     BC043617.1  GI:27694443
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 4294)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
```

```
REFERENCE   2  (bases 1 to 4294)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 89 Row: o Column: 19
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES              Location/Qualifiers
     source           1. .4294
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="MGC:50948 IMAGE:6150112"
                      /tissue_type="Skin, melanotic melanoma."
                      /clone_lib="NIH_MGC_72"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6"
     gene             1. .4294
                      /gene="DNMT3A"
                      /note="synonyms: DNMT3A2, M.HsaIIIA"
                      /db_xref="GeneID:1788"
                      /db_xref="MIM:602769"
     CDS              238. .2976
                      /gene="DNMT3A"
                      /codon_start=1
                      /product="DNA cytosine methyltransferase 3 alpha, isoform
                      a"
                      /protein_id="AAH43617.1"
                      /db_xref="GI:27694444"
                      /db_xref="GeneID:1788"
                      /db_xref="MIM:602769"
                      /translation="MPAMPSSGPGDTSSSAAEREEDRKDGEEQEEPRGKEERQEPSTT
                      ARKVGRPGRKRKHPPVESGDTPKDPAVISKSPSMAQDSGASELLPNGDLEKRSEPQPE
                      EGSPAGGQKGGAPAEGEGAAETLPEASRAVENGCCTPKEGRGAPAEAGKEQKETNIES
                      MKMEGSRGRLRGGLGWESSLRQRPMPRLTFQAGDPYYISKRKRDEWLARWKREAEKKA
                      KVIAGMNAVEENQGPGESQKVEEASPPAVQQPTDPASPTVATTPEPVGSDAGDKNATK
                      AGDDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDG
                      KFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPVCHDSDE
                      SDTAKAVEVQNKPMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYAP
                      PPPAKKPRKSTAEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPL
                      FVGGMCQNCKNCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLL
                      VGPGAAQAAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYP
                      PVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIM
                      YVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHD
                      ARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNL
                      PGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKE
                      DILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              0              Length:          4294
Score:                  3763.00        Matches:         689
Percent Similarity:     100.0%         Conservative:    0
Best Local Similarity:  100.0%         Mismatches:      0
Query Match:            100.0%         Indels:          0
DB:                     5              Gaps:            0

US-10-623-813-86 (1-689) x BC043617 (1-4294)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 966

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1026

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1086

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1146

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1206

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1266

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1326

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1386

Qy        161 ProProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1387 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1446

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1447 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1506

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1507 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1566

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1567 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1626

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1627 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1686

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1687 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1746

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1747 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1806

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1807 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1866
```

```
Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1867 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1926

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1927 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1986

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1987 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2046

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2047 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2106

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2107 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2166

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2167 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2226

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2227 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2286

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2287 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2346

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2347 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2406

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2407 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2466

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2467 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2526

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2527 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2586

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2587 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2646

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2647 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2706

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2707 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2766

Qy        621 LysAspGlnHisProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2767 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2826

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2827 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2886

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2887 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2946
```

```
Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              |||||||||||||||||||||||||||
Db       2947 CCGCTGAAGGAGTATTTTGCGTGTGTG 2973

RESULT 8
CS050241
LOCUS       CS050241                4395 bp    DNA     linear   PAT 23-MAR-2005
DEFINITION  Sequence 25 from Patent WO2005021757.
ACCESSION   CS050241
VERSION     CS050241.1  GI:61889496
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tomme,P.H. and van Rompaey,L.
  TITLE     Polypeptides and polynucleotides for use as a medicament
  JOURNAL   Patent: WO 2005021757-A 25 10-MAR-2005;
            Galapagos Genomics N.V. (BE)
FEATURES             Location/Qualifiers
     source          1. .4395
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          4395
Score:                    3763.00        Matches:         689
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       2              Gaps:            0

US-10-623-813-86 (1-689) x CS050241 (1-4395)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 1067

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1127

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1128 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1187

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1188 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1247

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1248 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1307

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1308 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1367

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1368 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1427

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1428 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1487

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1488 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1547
```

```
Qy       181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db      1548 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1607

Qy       201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db      1608 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1667

Qy       221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
Db      1668 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1727

Qy       241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db      1728 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1787

Qy       261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db      1788 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1847

Qy       281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db      1848 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1907

Qy       301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1908 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1967

Qy       321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db      1968 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 2027

Qy       341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
Db      2028 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 2087

Qy       361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
Db      2088 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2147

Qy       381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db      2148 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2207

Qy       401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      2208 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2267

Qy       421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      2268 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2327

Qy       441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      2328 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2387

Qy       461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db      2388 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2447

Qy       481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db      2448 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2507

Qy       501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      2508 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2567

Qy       521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
Db      2568 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2627

Qy       541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db      2628 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2687

Qy       561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db      2688 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2747

Qy       581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
Db      2748 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2807

Qy       601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db      2808 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2867

Qy       621 LysAspGlnHisProProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db      2868 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2927

Qy       641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2928 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2987

Qy       661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db      2988 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 3047

Qy       681 ProLeuLysGluTyrPheAlaCysVal 689
Db      3048 CCGCTGAAGGAGTATTTTGCGTGTGTG 3074
```

```
RESULT 9
CQ720469
LOCUS       CQ720469                2841 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 6403 from Patent WO02068579.
ACCESSION   CQ720469
VERSION     CQ720469.1  GI:42281326
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 6403 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .2841
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          2841
Score:                    3746.00        Matches:         689
Percent Similarity:       98.0%          Conservative:    0
Best Local Similarity:    98.0%          Mismatches:      0
Query Match:              99.5%          Indels:          14
DB:                       2              Gaps:            1

US-10-623-813-86 (1-689) x CQ720469 (1-2841)
```

```
Qy         1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
Db       730 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 789

Qy        21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
```

```
Db        790 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 849

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60

Db        850 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 909

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80

Db        910 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 969

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100

Db        970 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1029

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120

Db       1030 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1089

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140

Db       1090 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1149

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160

Db       1150 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1209

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180

Db       1210 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1269

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200

Db       1270 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1329

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220

Db       1330 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1389

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240

Db       1390 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1449

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260

Db       1450 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1509

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280

Db       1510 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1569

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300

Db       1570 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1629

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320

Db       1630 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1689

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340

Db       1690 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1749

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360

Db       1750 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1809

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380

Db       1810 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 1869

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400

Db       1870 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 1929

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420

Db       1930 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 1989

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440

Db       1990 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2049

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460

Db       2050 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2109

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480

Db       2110 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2169

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500

Db       2170 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2229

Qy        501 Tyr----------------------------------------GluGlyThrGlyArg 506

Db       2230 TACGGTAGACAGCCCCAGCTGATGGCTTTCTCTTCCGACCTCTCAGAGGGCACTGGCCGG 2289

Qy        507 LeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgProLysGluGlyAspAspArg 526

Db       2290 CTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCCAAGGAGGGAGATGATCGC 2349

Qy        527 ProPhePheTrpLeuPheGluAsnValValAlaMetGlyValSerAspLysArgAspIle 546

Db       2350 CCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTTAGTGACAAGAGGGACATC 2409

Qy        547 SerArgPheLeuGluSerAsnProValMetIleAspAlaLysGluValSerAlaAlaHis 566

Db       2410 TCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAAGAAGTGTCAGCTGCACAC 2469

Qy        567 ArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArgProLeuAlaSerThrVal 586

Db       2470 AGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGGCCGTTGGCATCCACTGTG 2529

Qy        587 AsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArgIleAlaLysPheSerLys 606

Db       2530 AATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGGATAGCCAAGTTCAGCAAA 2589

Qy        607 ValArgThrIleThrThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisPhePro 626

Db       2590 GTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGCAAAGACCAGCATTTTCCT 2649

Qy        627 ValPheMetAsnGluLysGluAspIleLeuTrpCysThrGluMetGluArgValPheGly 646

Db       2650 GTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAAATGGAAAGGGTATTTGGT 2709

Qy        647 PheProValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuLeu 666

Db       2710 TTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTGGCGAGGCAGAGACTGCTG 2769

Qy        667 GlyArgSerTrpSerValProValIleArgHisLeuPheAlaProLeuLysGluTyrPhe 686

Db       2770 GGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCTCCGCTGAAGGAGTATTTT 2829

Qy        687 AlaCysVal 689

Db       2830 GCGTGTGTG 2838
```

RESULT 10
AF480164
LOCUS AF480164 2318 bp mRNA linear ROD 14-OCT-2002
DEFINITION Mus musculus DNA cytosine methyltransferase 3A2 (Dnmt3a2) mRNA, complete cds.
ACCESSION AF480164
VERSION AF480164.1 GI:23954439

```
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2318)
  AUTHORS   Chen,T., Ueda,Y., Xie,S. and Li,E.
  TITLE     A Novel Dnmt3a Isoform Produced from an Alternative Promoter
            Localizes to Euchromatin and Its Expression Correlates with Active
            de Novo Methylation
  JOURNAL   J. Biol. Chem. 277 (41), 38746-38754 (2002)
   PUBMED   12138111
REFERENCE   2  (bases 1 to 2318)
  AUTHORS   Chen,T. and Li,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2002) Cardiovascular Research Center,
            Massachusetts General Hospital, 149 13th Street, Charlestown, MA
            02129, USA
FEATURES             Location/Qualifiers
     source          1. .2318
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129S4/SvJae"
                     /db_xref="taxon:10090"
                     /chromosome="12"
                     /map="12A2-A3"
     gene            1. .2318
                     /gene="Dnmt3a2"
     CDS             148. .2217
                     /gene="Dnmt3a2"
                     /function="de novo DNA cytosine methylation"
                     /codon_start=1
                     /product="DNA cytosine methyltransferase 3A2"
                     /protein_id="AAN40038.1"
                     /db_xref="GI:23954440"
                     /translation="MNAVEENQASGESQKVEEASPPAVQQPTDPASPTVATTPEPVGG
                     DAGDKNATKAADDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGT
                     RWVMWFGDGKFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKL
                     FPACHDSDESDSGKAVEVQNKQMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWV
                     EPEAAAYAPPPPAKKPRKSTTEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGS
                     LNVTLEHPLFIGGMCQNCKNCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRC
                     FCVECVDLLVGPGAAQAAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQ
                     EFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVG
                     MVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLF
                     FEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAH
                     RARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQH
                     FPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPL
                     KEYFACV"
ORIGIN

Alignment Scores:
Pred. No.:              0           Length:         2318
Score:                  3714.00     Matches:        679
Percent Similarity:     98.8%       Conservative:   2
Best Local Similarity:  98.5%       Mismatches:     8
Query Match:            98.7%       Indels:         0
DB:                     6           Gaps:           0

US-10-623-813-86 (1-689) x AF480164 (1-2318)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              |||||||||||||||||||||||      ||||||||||||||||||||||||||||||
Db        148 ATGAATGCTGTGGAAGAGAACCAGGCCTCTGGAGAGTCTCAGAAGGTGGAGGAGGCCAGC 207

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 CCTCCTGCTGTGCAGCAGCCCACGGACCCTGCTTCTCCGACTGTGGCCACCACCCCTGAG 267

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              |||||||||   ||||||||||||||||||||||||||||||   ||||||||||||||
Db        268 CCAGTAGGAGGGGATGCTGGGGACAAGAATGCTACCAAAGCAGCCGACGATGAGCCTGAG 327

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TATGAGGATGGCCGGGGCTTTGGCATTGGAGAGCTGGTGTGGGGGAAACTTCGGGGCTTC 387

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TCCTGGTGGCCAGGCCGAATTGTGTCTTGGTGGATGACAGGCCGGAGCCGAGCAGCTGAA 447

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 GGCACTCGCTGGGTCATGTGGTTCGGAGATGGCAAGTTCTCAGTGGTGTGTGTGGAGAAG 507

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 CTCATGCCGCTGAGCTCCTTCTGCAGTGCATTCCACCAGGCCACCTACAACAAGCAGCCC 567

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 ATGTACCGCAAAGCCATCTACGAAGTCCTCCAGGTGGCCAGCAGCCGTGCCGGGAAGCTG 627

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||   ||||||||||||||||||||||:::   ||||||||||||||||||||
Db        628 TTTCCAGCTTGCCATGACAGTGATGAAAGTGACAGTGGCAAGGCTGTGGAAGTGCAGAAC 687

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 AAGCAGATGATTGAATGGGCCCTCGGTGGCTTCCAGCCCTCGGGTCCTAAGGGCCTGGAG 747

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACCAGAAGAAGAGAAGAATCCTTACAAGGAAGTTTACACCGACATGTGGGTGGAGCCT 807

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        808 GAAGCAGCTGCTTACGCCCCACCCCCACCAGCCAAGAAACCCAGAAAGAGCACAACAGAG 867

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 AAACCTAAGGTCAAGGAGATCATTGATGAGCGCACAAGGGAGCGGCTGGTGTATGAGGTG 927

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 CGCCAGAAGTGCAGAAACATCGAGGACATTTGTATCTCATGTGGGAGCCTCAATGTCACC 987

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        988 CTGGAGCACCCACTCTTCATTGGAGGCATGTGCCAGAACTGTAAGAACTGCTTCTTGGAG 1047

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 TGTGCTTACCAGTATGACGACGATGGGTACCAGTCCTATTGCACCATCTGCTGTGGGGGG 1107

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 CGTGAAGTGCTCATGTGTGGGAACAACAACTGCTGCAGGTGCTTTTGTGTCGAGTGTGTG 1167

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 GATCTCTTGGTGGGGCCAGGAGCTGCTCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1227

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 TACATGTGCGGGCATAAGGGCACCTATGGGCTGCTGCGAAGACGGGAAGACTGGCCTTCT 1287

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 CGACTCCAGATGTTCTTTGCCAATAACCATGACCAGGAATTTGACCCCCCAAAGGTTTAC 1347

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CCACCTGTGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTCTTTGATGGGATT 1407

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTGGACCGCTACATTGCCTCC 1467

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGAAAGATCATGTAC 1527

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1528 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGACCTG 1587

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1588 GTGATTGGAGGCAGTCCCTGCAATGACCTCTCCATTGTCAACCCTGCCCGCAAGGGACTT 1647

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1648 TATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 1707

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1708 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 1767

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1768 AGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTGATGATTGACGCCAAA 1827

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1828 GAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAACAGG 1887

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1888 CCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCAGA 1947

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1948 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACTCTATAAAGCAGGGC 2007

Qy        621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2008 AAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATCCTGTGGTGCACTGAA 2067

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2068 ATGGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAACATGAGCCGCTTG 2127

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2128 GCGAGGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCT 2187

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
                ||||||||||||||||||||||||||
Db       2188 CCGCTGAAGGAATATTTTGCTTGTGTG 2214

RESULT 11
BC007466
LOCUS       BC007466               4094 bp    mRNA    linear   ROD 29-JUN-2004
DEFINITION  Mus musculus DNA methyltransferase 3A, transcript variant 1, mRNA
            (cDNA clone MGC:5662 IMAGE:3492853), complete cds.
ACCESSION   BC007466
VERSION     BC007466.1  GI:13938620
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4094)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 4094)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 8 Row: n Column: 11
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 6681208.
FEATURES             Location/Qualifiers
     source          1. .4094
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="Czech II"
                     /db_xref="taxon:10090"
                     /clone="MGC:5662 IMAGE:3492853"
                     /tissue_type="Mammary tumor metastatized to lung. Tumor
                     arose spontaneously from a senescent normal mammary
                     (clonal) outgrowth infected with the virus MMTV."
                     /clone_lib="NCI_CGAP_Lu29"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .4094
                     /gene="Dnmt3a"
                     /db_xref="GeneID:13435"
                     /db_xref="MGI:1261827"
     CDS             245. .2971
                     /gene="Dnmt3a"
                     /codon_start=1
                     /product="DNA methyltransferase 3A, isoform 1"
                     /protein_id="AAH07466.1"
                     /db_xref="GI:13938621"
                     /db_xref="GeneID:13435"
                     /db_xref="MGI:1261827"
                     /translation="MPSSGPGDTSSSSLEREDDRKEGEEQEENRGKEERQEPSATARK
                     VGRPGRKRKHPPVESSDTPKDPAVTTKSQPMAQDSGPSDLLPNGDLEKRSEPQPEEGS
                     PAAGQKGGAPAEGEGTETPPEASRAVENGCCVTKEGRGASAGEGKEQKPTNIESMKME
                     GSRGRLRGGLGWESSLRQRPMPRLTFQAGDPYYISKRKRDEWLARWKREAEKKAKVIA
                     VMNAVEENQASGESQKVEEASPPAVQQPTDPASPTVATTPEPVGGDAGDKNATKAADD
```

EPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSV
VCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPACHDSDESDSG
KAVEVQNKQMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYAPPPPA
KKPRKSTTEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPLFIGG
MCQNCKNCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLLVGPG
AAQAAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYPPVPA
EKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGD
VRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPK
EGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMN
RPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILW
CTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	4094
Score:	3714.00	Matches:	679
Percent Similarity:	98.8%	Conservative:	2
Best Local Similarity:	98.5%	Mismatches:	8
Query Match:	98.7%	Indels:	0
DB:	6	Gaps:	0

US-10-623-813-86 (1-689) x BC007466 (1-4094)

```
Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||      ||||||||||||||||||||||||||||||
Db        902 ATGAATGCTGTGGAAGAGAACCAGGCCTCTGGAGAGTCTCAGAAGGTGGAGGAGGCCAGC 961

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 CCTCCTGCTGTGCAGCAGCCCACGGACCCTGCTTCTCCGACTGTGGCCACCACCCCTGAG 1021

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              |||||||||  ||||||||||||||||||||||||||||||  ||||||||||||||||
Db       1022 CCAGTAGGAGGGGATGCTGGGGACAAGAATGCTACCAAAGCAGCCGACGATGAGCCTGAG 1081

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 TATGAGGATGGCCGGGGCTTTGGCATTGGAGAGCTGGTGTGGGGGAAACTTCGGGGCTTC 1141

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 TCCTGGTGGCCAGGCCGAATTGTGTCTTGGTGGATGACAGGCCGGAGCCGAGCAGCTGAA 1201

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 GGCACTCGCTGGGTCATGTGGTTCGGAGATGGCAAGTTCTCAGTGGTGTGTGTGGAGAAG 1261

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 CTCATGCCGCTGAGCTCCTTCTGCAGTGCATTCCACCAGGCCACCTACAACAAGCAGCCC 1321

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 ATGTACCGCAAAGCCATCTACGAAGTCCTCCAGGTGGCCAGCAGCCGTGCCGGGAAGCTG 1381

Qy        161 ProProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||  ||||||||||||||||||||||:::  ||||||||||||||||||||||||
Db       1382 TTTCCAGCTTGCCATGACAGTGATGAAAGTGACAGTGGCAAGGCTGTGGAAGTGCAGAAC 1441

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 AAGCAGATGATTGAATGGGCCCTCGGTGGCTTCCAGCCCTCGGGTCCTAAGGGCCTGGAA 1501

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 CCACCAGAAGAAGAGAAGAATCCTTACAAGGAAGTTTACACCGACATGTGGGTGGAGCCT 1561

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||
Db       1562 GAAGCAGCTGCTTACGCCCCACCCCCACCAGCCAAGAAACCCAGAAAGAGCACAACAGAG 1621

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1622 AAACCTAAGGTCAAGGAGATCATTGATGAGCGCACAAGGGAGCGGCTGGTGTATGAGGTG 1681

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 CGCCAGAAGTGCAGAAACATCGAGGACATTTGTATCTCATGTGGGAGCCTCAATGTCACC 1741

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||
Db       1742 CTGGAGCACCCACTCTTCATTGGAGGCATGTGCCAGAACTGTAAGAACTGCTTCTTGGAG 1801

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 TGTGCTTACCAGTATGACGACGATGGGTACCAGTCCTATTGCACCATCTGCTGTGGGGGG 1861

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 CGTGAAGTGCTCATGTGTGGGAACAACAACTGCTGCAGGTGCTTTTGTGTCGAGTGTGTG 1921

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 GATCTCTTGGTGGGGCCAGGAGCTGCTCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1981

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 TACATGTGCGGGCATAAGGGCACCTATGGGCTGCTGCGAAGACGGGAAGACTGGCCTTCT 2041

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 CGACTCCAGATGTTCTTTGCCAATAACCATGACCAGGAATTTGACCCCCCAAAGGTTTAC 2101

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 CCACCTGTGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTCTTTGATGGGATT 2161

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2162 GCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTGGACCGCTACATTGCCTCC 2221

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2222 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGAAAGATCATGTAC 2281

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2282 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGACCTG 2341

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2342 GTGATTGGAGGCAGTCCCTGCAATGACCTCTCCATTGTCAACCCTGCCCGCAAGGGACTT 2401

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2402 TATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2461

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2462 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2521

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2522 AGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTGATGATTGACGCCAAA 2581

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2582 GAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAACAGG 2641

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2642 CCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCAGA 2701

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2702 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACTCTATAAAGCAGGGC 2761
```

```
Qy        621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2762 AAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATCCTGTGGTGCACTGAA 2821

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2822 ATGGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAACATGAGCCGCTTG 2881

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2882 GCGAGGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCT 2941

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              |||||||||||||||||||||||||||
Db       2942 CCGCTGAAGGAATATTTTGCTTGTGTG 2968
```

```
RESULT 12
AF068625
LOCUS       AF068625               4192 bp    mRNA    linear   ROD 06-DEC-1999
DEFINITION  Mus musculus DNA cytosine-5 methyltransferase 3A (Dnmt3a) mRNA,
            complete cds.
ACCESSION   AF068625
VERSION     AF068625.2  GI:6449467
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4192)
  AUTHORS   Okano,M., Xie,S. and Li,E.
  TITLE     Cloning and characterization of a family of novel mammalian DNA
            (cytosine-5) methyltransferases
  JOURNAL   Nat. Genet. 19 (3), 219-220 (1998)
   PUBMED   9662389
REFERENCE   2  (bases 1 to 4192)
  AUTHORS   Xie,S., Okano,M. and Li,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAY-1998) CVRC, Mass. Gen. Hospital, 149 13th Street,
            Charlestown, MA 02129, USA
REFERENCE   3  (bases 1 to 4192)
  AUTHORS   Okano,M., Chijiwa,T., Sasaki,H. and Li,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-NOV-1999) CVRC, Mass. Gen. Hospital, 149 13th Street,
            Charlestown, MA 02129, USA
  REMARK    Sequence update by submitter
COMMENT     On Nov 18, 1999 this sequence version replaced gi:3327977.
FEATURES             Location/Qualifiers
     source          1. .4192
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /chromosome="12"
                     /map="4.0 cM"
     gene            1. .4192
                     /gene="Dnmt3a"
     misc_feature    1. .4192
                     /gene="Dnmt3a"
                     /note="similar to EST sequences deposited in GenBank
                     Accession Numbers AA052791, AA111043, AA154890, AA240794,
                     AA756653, W58898, W59299, W91664, and W91665"
     CDS             217. .2943
                     /gene="Dnmt3a"
                     /function="de novo DNA methylation"
                     /note="contains Cys-rich region"
                     /codon_start=1
                     /product="DNA cytosine-5 methyltransferase 3A"
                     /protein_id="AAC40177.2"
                     /db_xref="GI:6449468"
                     /translation="MPSSGPGDTSSSSLEREDDRKEGEEQEENRGKEERQEPSATARK
                     VGRPGRKRKHPPVESSDTPKDPAVTTKSQPMAQDSGPSDLLPNGDLEKRSEPQPEEGS
                     PAAGQKGGAPAEGEGTETPPEASRAVENGCCVTKEGRGASAGEGKEQKQTNIESMKME
                     GSRGRLRGGLGWESSLRQRPMPRLTFQAGDPYYISKRKRDEWLARWKREAEKKAKVIA
                     VMNAVEENQASGESQKVEEASPPAVQQPTDPASPTVATTPEPVGGDAGDKNATKAADD
                     EPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSV
                     VCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPACHDSDESDSG
                     KAVEVQNKQMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYAPPPPA
                     KKPRKSTTEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPLFIGG
                     MCQNCKNCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLLVGPG
                     AAQAAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYPPVPA
                     EKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGD
                     VRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPK
                     EGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMN
                     RPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILW
                     CTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV"
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          4192
Score:                    3714.00        Matches:         679
Percent Similarity:       98.8%          Conservative:    2
Best Local Similarity:    98.5%          Mismatches:      8
Query Match:              98.7%          Indels:          0
DB:                       6              Gaps:            0

US-10-623-813-86 (1-689) x AF068625 (1-4192)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||      ||||||||||||||||||||||||||||||
Db        874 ATGAATGCTGTGGAAGAGAACCAGGCCTCTGGAGAGTCTCAGAAGGTGGAGGAGGCCAGC 933

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        934 CCTCCTGCTGTGCAGCAGCCCACGGACCCTGCTTCTCCGACTGTGGCCACCACCCCTGAG 993

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              |||||||||  ||||||||||||||||||||||||||||||    ||||||||||||||||
Db        994 CCAGTAGGAGGGGATGCTGGGGACAAGAATGCTACCAAAGCAGCCGACGATGAGCCTGAG 1053

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1054 TATGAGGATGGCCGGGGCTTTGGCATTGGAGAGCTGGTGTGGGGGAAACTTCGGGGCTTC 1113

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1114 TCCTGGTGGCCAGGCCGAATTGTGTCTTGGTGGATGACAGGCCGGAGCCGAGCAGCTGAA 1173

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1174 GGCACTCGCTGGGTCATGTGGTTCGGAGATGGCAAGTTCTCAGTGGTGTGTGTGGAGAAG 1233

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1234 CTCATGCCGCTGAGCTCCTTCTGCAGTGCATTCCACCAGGCCACCTACAACAAGCAGCCC 1293

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1294 ATGTACCGCAAAGCCATCTACGAAGTCCTCCAGGTGGCCAGCAGCCGTGCCGGGAAGCTG 1353

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||   ||||||||||||||||||||||:::  ||||||||||||||||||||||
Db       1354 TTTCCAGCTTGCCATGACAGTGATGAAAGTGACAGTGGCAAGGCTGTGGAAGTGCAGAAC 1413

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1414 AAGCAGATGATTGAATGGGCCCTCGGTGGCTTCCAGCCCTCGGGTCCTAAGGGCCTGGAG 1473

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1474 CCACCAGAAGAAGAGAAGAATCCTTACAAGGAAGTTTACACCGACATGTGGGTGGAGCCT 1533

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||    |||
Db       1534 GAAGCAGCTGCTTACGCCCCACCCCCACCAGCCAAGAAACCCAGAAAGAGCACAACAGAG 1593
```

```
Qy       241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db      1594 AAACCTAAGGTCAAGGAGATCATTGATGAGCGCACAAGGGAGCGGCTGGTGTATGAGGTG 1653

Qy       261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db      1654 CGCCAGAAGTGCAGAAACATCGAGGACATTTGTATCTCATGTGGGAGCCTCAATGTCACC 1713

Qy       281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db      1714 CTGGAGCACCCACTCTTCATTGGAGGCATGTGCCAGAACTGTAAGAACTGCTTCTTGGAG 1773

Qy       301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1774 TGTGCTTACCAGTATGACGACGATGGGTACCAGTCCTATTGCACCATCTGCTGTGGGGGG 1833

Qy       321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db      1834 CGTGAAGTGCTCATGTGTGGGAACAACAACTGCTGCAGGTGCTTTTGTGTCGAGTGTGTG 1893

Qy       341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
Db      1894 GATCTCTTGGTGGGGCCAGGAGCTGCTCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1953

Qy       361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
Db      1954 TACATGTGCGGGCATAAGGGCACCTATGGGCTGCTGCGAAGACGGGAAGACTGGCCTTCT 2013

Qy       381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db      2014 CGACTCCAGATGTTCTTTGCCAATAACCATGACCAGGAATTTGACCCCCCAAAGGTTTAC 2073

Qy       401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      2074 CCACCTGTGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTCTTTGATGGGATT 2133

Qy       421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      2134 GCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTGGACCGCTACATTGCCTCC 2193

Qy       441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      2194 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGAAAGATCATGTAC 2253

Qy       461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db      2254 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGACCTG 2313

Qy       481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db      2314 GTGATTGGAGGCAGTCCCTGCAATGACCTCTCCATTGTCAACCCTGCCCGCAAGGGACTT 2373

Qy       501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      2374 TATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2433

Qy       521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
Db      2434 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2493

Qy       541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db      2494 AGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTGATGATTGACGCCAAA 2553

Qy       561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db      2554 GAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAACAGG 2613

Qy       581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
Db      2614 CCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCAGA 2673

Qy       601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db      2674 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACTCTATAAAGCAGGGC 2733

Qy       621 LysAspGlnHisProProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db      2734 AAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATCCTGTGGTGCACTGAA 2793

Qy       641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2794 ATGGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAACATGAGCCGCTTG 2853

Qy       661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db      2854 GCGAGGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCT 2913

Qy       681 ProLeuLysGluTyrPheAlaCysVal 689
Db      2914 CCGCTGAAGGAATATTTTGCTTGTGTG 2940
```

RESULT 13
AY271299

```
LOCUS       AY271299               2597 bp    mRNA    linear   MAM 05-SEP-2003
DEFINITION  Bos taurus DNA methyltransferase 3a isoform 4 (DNMT3A) mRNA,
            partial cds.
ACCESSION   AY271299
VERSION     AY271299.1  GI:32350982
KEYWORDS    .
SOURCE      Bos taurus (cattle)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 2597)
  AUTHORS   Golding,M.C. and Westhusin,M.E.
  TITLE     Analysis of DNA (cytosine 5) Methyltransferase mRNA Sequence and
            Expression in Bovine Preimplantation Embryos, Fetal and Adult
            Tissues
  JOURNAL   Gene Expr. Patterns 3 (5), 551-558 (2003)
REFERENCE   2  (bases 1 to 2597)
  AUTHORS   Golding,M.C. and Westhusin,M.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-APR-2003) Veterinary Physiology, Texas A&M - CVM, 500
            University Dr. West, College Station, TX 77843-4466, USA
FEATURES             Location/Qualifiers
     source          1. .2597
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
     gene            <1. .2597
                     /gene="DNMT3A"
     CDS             <1. .2467
                     /gene="DNMT3A"
                     /note="DNMT3a4; DNA (cytosine 5) methyltransferase"
                     /codon_start=2
                     /product="DNA methyltransferase 3a isoform 4"
                     /protein_id="AAP75902.1"
                     /db_xref="GI:32350983"
                     /translation="EGEEQEEARAKEERQEPSTTARKVGRPGRKRKHPPVESSDTPKD
                     PAVTSKSLSMAQDSGPSELYPNGDLEKRSEPQPEEGSPAGGQKGGAPAEGEGATETPP
                     EASRAVENGCCTPKEPLQKKAEKKAKVIAVMNAVEENQGSTESQKVEEASPPAVQQPT
                     DPASPTVATTPEPVGADAGDKNATKAADDEPEYEDGRGFGIGELVWGKLRGFSWWPGR
                     IVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRK
                     AIYEVLQVASSRAGKLFPMCHDSDESDTAKAVEVQNKQMIEWALGGFQPSGPKGLEPP
                     EEEKNPYKEVYTDMWVEPEAAAYAPPPPAKKPRKSTTEKPKVKEIIDERTRERLVYEV
                     RQKCRNIEDICISCGSLNVTLEHPLFIGGMCQNCKNCFLECAYQYDDDGYQSYCTICC
                     GGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIKEDPWNCYMCGHKGTYGLLRRRD
                     DWPSRLQMFFANNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQV
                     DRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSI
                     VNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVVLGVSDKRDISRFL
                     ESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVR
                     TITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLL
                     GRSWSVPVIRHLFAPLKEYFACV"
ORIGIN
```

Alignment Scores:

Pred. No.:	0	Length:	2597
Score:	3712.00	Matches:	678
Percent Similarity:	99.1%	Conservative:	5
Best Local Similarity:	98.4%	Mismatches:	6
Query Match:	98.6%	Indels:	0
DB:	14	Gaps:	0

US-10-623-813-86 (1-689) x AY271299 (1-2597)

```
Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||      ||||||||||||||||||||||||||||||
Db        398 ATGAATGCTGTTGAAGAAAACCAGGGGTCCACCGAGTCTCAGAAGGTGGAGGAGGCCAGT 457

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 CCTCCCGCTGTGCAGCAGCCCACCGACCCTGCGTCCCCCACAGTGGCCACCACGCCTGAG 517

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||:::||||||||||||||||||||||||||||||      ||||||||||||||||
Db        518 CCCGTGGGCGCCGATGCCGGGGACAAGAATGCCACCAAAGCAGCTGACGATGAACCGGAG 577

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 637

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCGGAA 697

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 GGCACTCGTTGGGTCATGTGGTTCGGAGACGGCAAGTTCTCAGTGGTGTGTGTGGAGAAG 757

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 CTGATGCCGCTGAGCTCCTTCTGCAGTGCTTTCCACCAGGCCACCTACAACAAGCAGCCC 817

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 ATGTACCGCAAGGCTATCTACGAGGTCCTGCAGGTGGCCAGCAGTCGAGCGGGGAAGCTG 877

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 TTTCCAATGTGCCATGACAGCGACGAGAGCGACACTGCCAAGGCCGTGGAGGTGCAGAAC 937

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 AAACAAATGATCGAGTGGGCCCTGGGAGGGTTCCAGCCCTCTGGCCCCAAGGGCCTGGAG 997

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 CCCCCAGAAGAGGAGAAGAACCCCTACAAAGAAGTTTACACAGACATGTGGGTTGAACCC 1057

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db       1058 GAGGCAGCTGCCTATGCGCCGCCCCCACCAGCCAAAAAGCCCCGAAAGAGCACAACGGAG 1117

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1118 AAGCCTAAGGTCAAGGAGATCATTGATGAACGCACAAGAGAGCGGCTGGTGTACGAGGTA 1177

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1178 CGGCAGAAGTGCCGGAACATCGAGGACATCTGCATCTCTTGTGGGAGCCTCAACGTCACC 1237

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db       1238 TTGGAACACCCTCTCTTCATCGGAGGAATGTGCCAAAACTGCAAGAACTGCTTCCTGGAA 1297

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1298 TGCGCGTACCAGTACGATGACGATGGCTATCAGTCCTACTGCACCATCTGCTGCGGGGGG 1357

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1358 CGCGAGGTGCTCATGTGTGGGAACAACAATTGCTGCAGGTGCTTTTGCGTGGAATGTGTG 1417

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1418 GATCTCTTGGTGGGGCCAGGGGCCGCGCAGGCAGCCATCAAAGAAGACCCCTGGAACTGC 1477

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||
Db       1478 TACATGTGCGGGCACAAGGGCACCTACGGGCTGCTGCGGCGGCGGGACGACTGGCCGTCT 1537

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1538 CGGCTCCAGATGTTCTTCGCCAACAACCATGACCAGGAATTCGATCCTCCGAAGGTTTAC 1597

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1598 CCACCTGTCCCAGCTGAGAAGAGAAAGCCCATCCGGGTGCTGTCTCTATTCGATGGAATT 1657

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1658 GCTACAGGGCTTCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATCGCCTCC 1717

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1718 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGAAAGATCATGTAC 1777

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1778 GTCGGGGACGTCCGCAGCGTTACACAGAAGCATATCCAGGAATGGGGCCCGTTCGATCTG 1837

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1838 GTGATTGGGGGCAGTCCCTGCAATGATCTCTCCATCGTCAACCCTGCCCGCAAGGGACTC 1897

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1898 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 1957

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   :::||||||
Db       1958 AAGGAGGGAGATGACCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGTCCTGGGCGTT 2017

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2018 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2077

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2078 GAAGTGTCAGCTGCGCACAGGGCCCGCTACTTCTGGGGGAACCTTCCTGGTATGAACAGG 2137

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2138 CCATTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCACGGCCGA 2197

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2198 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACTAGGTCGAACTCCATAAAGCAGGGC 2257

Qy        621 LysAspGlnHisProProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2258 AAAGACCAGCATTTCCCCGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2317

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2318 ATGGAAAGGGTGTTTGGCTTCCCTGTCCACTATACCGACGTCTCCAACATGAGCCGCTTG 2377

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2378 GCGAGGCAGAGACTGCTGGGCCGGTCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCT 2437
```

```
Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              |||||||||||||||||||||||||||||
Db       2438 CCGCTGAAGGAATATTTTGCTTGTGTG 2464

RESULT 14
AY271298
LOCUS       AY271298               2798 bp    mRNA    linear   MAM 05-SEP-2003
DEFINITION  Bos taurus DNA methyltransferase 3a (DNMT3A) mRNA, partial cds.
ACCESSION   AY271298
VERSION     AY271298.1  GI:32350980
KEYWORDS    .
SOURCE      Bos taurus (cattle)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 2798)
  AUTHORS   Golding,M.C. and Westhusin,M.E.
  TITLE     Analysis of DNA (cytosine 5) Methyltransferase mRNA Sequence and
            Expression in Bovine Preimplantation Embryos, Fetal and Adult
            Tissues
  JOURNAL   Gene Expr. Patterns 3 (5), 551-558 (2003)
REFERENCE   2  (bases 1 to 2798)
  AUTHORS   Golding,M.C. and Westhusin,M.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-APR-2003) Veterinary Physiology, Texas A&M - CVM, 500
            University Dr. West, College Station, TX 77843-4466, USA
FEATURES             Location/Qualifiers
     source          1. .2798
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
     gene            <1. .2798
                     /gene="DNMT3A"
     CDS             <1. .2668
                     /gene="DNMT3A"
                     /note="DNMT3a; DNA (cytosine 5) methyltransferase"
                     /codon_start=2
                     /product="DNA methyltransferase 3a"
                     /protein_id="AAP75901.1"
                     /db_xref="GI:32350981"
                     /translation="EGEEQEEARAKEERQEPSTTARKVGRPGRKRKHPPVESSDTPKD
                     PAVTSKSLSMAQDSGPSELYPNGDLEKRSEPQPEEGSPAGGQKGGAPAEGEGATETPP
                     EASRAVENGCCTPKDGRGAPAEEGKEQKETNIESMKMEGSRGRLRGGLGWESSLRQRP
                     MPRLTFQAGDPYYISKRKRDEWLARWKREAEKKAKVIAVMNAVEENQGSTESQKVEEA
                     SPPAVQQPTDPASPTVATTPEPVGADAGDKNATKAADDEPEYEDGRGFGIGELVWGKL
                     RGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCVEKLMPLSSFCSAFHQAT
                     YNKQPMYRKAIYEVLQVASSRAGKLFPMCHDSDESDTAKAVEVQNKQMIEWALGGFQP
                     SGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYAPPPPAKKPRKSTTEKPKVKEIIDER
                     TRERLVYEVRQKCRNIEDICISCGSLNVTLEHPLFIGGMCQNCKNCFLECAYQYDDDG
                     YQSYCTICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIKEDPWNCYMCGHKG
                     TYGLLRRRDDWPSRLQMFFANNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLL
                     VLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIG
                     GSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVVLGVS
                     DKRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHG
                     RIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNM
                     SRLARQRLLGRSWSVPVIRHLFAPLKEYFACV"
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:         2798
Score:                  3712.00        Matches:        678
Percent Similarity:     99.1%          Conservative:   5
Best Local Similarity:  98.4%          Mismatches:     6
Query Match:            98.6%          Indels:         0
DB:                     14             Gaps:           0

US-10-623-813-86 (1-689) x AY271298 (1-2798)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              |||||||||||||||||||||||||      |||||||||||||||||||||||||||||
Db        599 ATGAATGCTGTTGAAGAAAACCAGGGGTCCACCGAGTCTCAGAAGGTGGAGGAGGCCAGT 658

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CCTCCCGCTGTGCAGCAGCCCACCGACCCTGCGTCCCCCACAGTGGCCACCACGCCTGAG 718

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              |||||||||:::||||||||||||||||||||||||||||||    ||||||||||||||
Db        719 CCCGTGGGCGCCGATGCCGGGGACAAGAATGCCACCAAAGCAGCTGACGATGAACCGGAG 778

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 838

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCGGAA 898

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 GGCACTCGTTGGGTCATGTGGTTCGGAGACGGCAAGTTCTCAGTGGTGTGTGTGGAGAAG 958

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CTGATGCCGCTGAGCTCCTTCTGCAGTGCTTTCCACCAGGCCACCTACAACAAGCAGCCC 1018

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 ATGTACCGCAAGGCTATCTACGAGGTCCTGCAGGTGGCCAGCAGTCGAGCGGGGAAGCTG 1078

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 TTTCCAATGTGCCATGACAGCGACGAGAGCGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1138

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 AAACAAATGATCGAGTGGGCCCTGGGAGGGTTCCAGCCCTCTGGCCCCAAGGGCCTGGAG 1198

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 CCCCCAGAAGAGGAGAAGAACCCCTACAAAGAAGTTTACACAGACATGTGGGTTGAACCC 1258

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    |||||
Db       1259 GAGGCAGCTGCCTATGCGCCGCCCCCACCAGCCAAAAAGCCCCGAAAGAGCACAACGGAG 1318

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 AAGCCTAAGGTCAAGGAGATCATTGATGAACGCACAAGAGAGCGGCTGGTGTACGAGGTA 1378

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 CGGCAGAAGTGCCGGAACATCGAGGACATCTGCATCTCTTGTGGGAGCCTCAACGTCACC 1438

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db       1439 TTGGAACACCCTCTCTTCATCGGAGGAATGTGCCAAAACTGCAAGAACTGCTTCCTGGAA 1498

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 TGCGCGTACCAGTACGATGACGATGGCTATCAGTCCTACTGCACCATCTGCTGCGGGGGG 1558

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1559 CGCGAGGTGCTCATGTGTGGGAACAACAATTGCTGCAGGTGCTTTTGCGTGGAATGTGTG 1618

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1619 GATCTCTTGGTGGGGCCAGGGGCCGCGCAGGCAGCCATCAAAGAAGACCCCTGGAACTGC 1678

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db       1679 TACATGTGCGGGCACAAGGGCACCTACGGGCTGCTGCGGCGGCGGGACGACTGGCCGTCT 1738
```

```
Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1739 CGGCTCCAGATGTTCTTCGCCAACAACCATGACCAGGAATTCGATCCTCCGAAGGTTTAC 1798

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1799 CCACCTGTCCCAGCTGAGAAGAGAAAGCCCATCCGGGTGCTGTCTCTATTCGATGGAATT 1858

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1859 GCTACAGGGCTTCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATCGCCTCC 1918

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1919 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGAAAGATCATGTAC 1978

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1979 GTCGGGGACGTCCGCAGCGTTACACAGAAGCATATCCAGGAATGGGGCCCGTTCGATCTG 2038

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2039 GTGATTGGGGGCAGTCCCTGCAATGATCTCTCCATCGTCAACCCTGCCCGCAAGGGACTC 2098

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2099 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2158

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2159 AAGGAGGGAGATGACCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGTCCTGGGCGTT 2218

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2278

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2279 GAAGTGTCAGCTGCGCACAGGGCCCGCTACTTCTGGGGGAACCTTCCTGGTATGAACAGG 2338

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2339 CCATTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCACGGCCGA 2398

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2399 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACTAGGTCGAACTCCATAAAGCAGGGC 2458

Qy        621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2459 AAAGACCAGCATTTCCCCGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2518

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2519 ATGGAAAGGGTGTTTGGCTTCCCTGTCCACTATACCGACGTCTCCAACATGAGCCGCTTG 2578

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2579 GCGAGGCAGAGACTGCTGGGCCGGTCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCT 2638

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              ||||||||||||||||||||||||||
Db       2639 CCGCTGAAGGAATATTTTGCTTGTGTG 2665

RESULT 15
AB208833
LOCUS       AB208833               4476 bp    mRNA    linear   PRI 31-MAR-2005
DEFINITION  Homo sapiens mRNA for DNA cytosine methyltransferase 3 alpha
            isoform a variant protein.
ACCESSION   AB208833
VERSION     AB208833.1  GI:62087245
KEYWORDS    FLI_CDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y., Tanaka,A., Yokoyama,S.,
            Ohara,O., Nagase,T. and F.Kikuno,R.
  TITLE     None Title
  JOURNAL   Published Only in Database (2005)
REFERENCE   2  (bases 1 to 4476)
  AUTHORS   Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y., Tanaka,A., Yokoyama,S.,
            Ohara,O., Nagase,T. and F.Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2005) Osamu Ohara, Kazusa DNA Research Institute,
            Department of Human Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
            Chiba, 292-0818, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://protein.gsc.riken.go.jp/, Tel:81-438-52-3930,
            Fax:81-438-52-3931)
COMMENT     This work was supported in part by the National Project on Protein
            Structural and Functional Analysis , Ministry of Education, Culture
            , Sports, Science and Technology of Japan. Totoki Y, Toyoda A,
            Takeda T, Sakaki Y, Tanaka A, Yokoyama S. RIKEN Genomic Sciences
            Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan.
            e-mail: aktanaka@postman.riken.go.jp
            URL: http://protein.gsc.riken.go.jp/.
FEATURES            Location/Qualifiers
     source         1. .4476
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="pj01914"
                    /tissue_type="brain"
                    /note="this clone is also named as hsk003000130
                    vector:pBluescriptII SK plus"
     gene           <1. .4476
                    /gene="DNA cytosine methyltransferase 3 alpha isoform a
                    variant"
     CDS            <148. .2583
                    /gene="DNA cytosine methyltransferase 3 alpha isoform a
                    variant"
                    /inference="non-experimental evidence, no additional
                    details recorded"
                    /note="Start codon is not identified."
                    /codon_start=1
                    /product="DNA cytosine methyltransferase 3 alpha isoform a
                    variant"
                    /protein_id="BAD92070.1"
                    /db_xref="GI:62087246"
                    /translation="LSPPKTTIIPSPKPSSPPAPRSPSLPPGAQMPAMPSSGPGDTSS
                    SAAEREEDRKDGEEQEEPRGKEERQEPSTTARKVGRPGRKRKHPPVESGDTPKDPAVI
                    SKSPSMAQDSGASELLPNGDLEKRSEPQPEEGSPAGGQKGGAPAEGEGAAETLPEASR
                    AVENGCCTPKEGRGAPAEAGKEQKETNIESMKMEGSRGRLRGGLGWESSLRQRPMPRL
                    TFQAGDPYYISKRKRDEWLARWKREAEKKAKVIAGMNAVEENQGPGESQKVEEASPPA
                    VQQPTDPASPTVATTPEPVGSDAGDKNATKAGDDEPEYEDGRGFGIGELVWGKLRGFS
                    WWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCVEKLMPLSSFCSAFHQATYNKQ
                    PMYRKAIYEVLQVASSRAGKLFPVCHDSDESDTAKAVEVQNKPMIEWALGGFQPSGPK
                    GLEPPEEEKNPYKEVYTDMWVEPEAAAYAPPPPAKKPRKSTAEKPKVKEIIDERTRER
                    LVYEVRQKCRNIEDICISCGSLNVTLEHPLFVGGMCQNCKNCFLECAYQYDDDGYQSY
                    CTICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIKEDPWNCYMCGHKGTYGL
                    LRRREDWPSRLQMFFANNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKD
                    LGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPC
                    NDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRD
                    ISRFLEVQNSHLY"
ORIGIN

Alignment Scores:
Pred. No.:              0           Length:         4476
Score:                  3702.00     Matches:        689
Percent Similarity:     91.1%       Conservative:   0
Best Local Similarity:  91.1%       Mismatches:     0
Query Match:            98.4%       Indels:         67
DB:                     5           Gaps:           1
```

US-10-623-813-86 (1-689) x AB208833 (1-4476)

```
Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 966

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1026

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1027 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1086

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1087 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1146

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1147 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1206

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1207 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1266

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1267 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1326

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1327 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1386

Qy        161 ProProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1387 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1446

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1447 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1506

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1507 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1566

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1567 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1626

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1627 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1686

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1687 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1746

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1747 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1806

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1807 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1866

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1867 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1926

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1927 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1986

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1987 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2046

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2047 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2106

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2107 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2166

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2167 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2226

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2227 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2286

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2287 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2346

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2347 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2406

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2407 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2466

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2467 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2526

Qy        541 SerAspLysArgAspIleSerArgPheLeu---------------------------- 550
              |||||||||||||||||||||||||||||
Db       2527 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGGTACAAAATAGCCACCTGTACTAGTGG 2586

Qy        550 ------------------------------------------------------------ 550

Db       2587 GCTAGTGGCCACTGAATTGAACAGTGTAGAGGACAACCCACCATCGCAGAAGTTCATTTG 2646

Qy        550 ------------------------------------------------------------ 550

Db       2647 GACAATGCTGATCTAGAACGTTCCAGTGACAGCTTGTGGAATGTGGCTAGGTGTAATTCC 2706

Qy        551 ----------------------------------------------GluSerAsnPr 554
                                                             ||||||||||
Db       2707 AGCTTCTCCTGGAGGCTGCAGGCTAGCCCAGTGTGTGGCTCCTGAGAGAGAATCCAACCC 2766

Qy        554 oValMetIleAspAlaLysGluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAs 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2767 TGTGATGATTGATGCCAAAGAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAA 2826

Qy        574 nLeuProGlyMetAsnArgProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGl 594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2827 CCTTCCCGGTATGAACAGGCCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGA 2886

Qy        594 uCysLeuGluHisGlyArgIleAlaLysPheSerLysValArgThrIleThrThrArgSe 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2887 GTGTCTGGAGCATGGCAGGATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTC 2946

Qy        614 rAsnSerIleLysGlnGlyLysAspGlnHisProProValPheMetAsnGluLysGluAs 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2947 AAACTCCATAAAGCAGGGCAAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGA 3006

Qy        634 pIleLeuTrpCysThrGluMetGluArgValPheGlyPheProValHisTyrThrAspVa 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3007 CATCTTATGGTGCACTGAAATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGT 3066
```

```
Qy        654 lSerAsnMetSerArgLeuAlaArgGlnArgLeuLeuGlyArgSerTrpSerValProVa 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3067 CTCCAACATGAGCCGCTTGGCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGT 3126

Qy        674 lIleArgHisLeuPheAlaProLeuLysGluTyrPheAlaCysVal 689
              |||||||||||||||||||||||||||||||||||||||||||||
Db       3127 CATCCGCCACCTCTTCGCTCCGCTGAAGGAGTATTTTGCGTGTGTG 3172
```

Search completed: October 5, 2006, 21:09:00
Job time : 10579 secs

GenCore version 5.1.9

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 5, 2006, 21:00:34 ; Search time 2088 Seconds
(without alignments)
6082.026 Million cell updates/sec

Title: US-10-623-813-86
Perfect score: 3763
Sequence: 1 MNAVEENQGPGESQKVEEAS.....WSVPVIRHLFAPLKEYFACV 689

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10623813/runat_05102006_121153_27814/app_query.fasta_1
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10623813_@CGN_1_1_1675_@runat_05102006_121153_27814 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_Main:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3763	100.0	2371	9	US-10-623-813-84	Sequence 84, Appl
2	3763	100.0	3005	6	US-10-144-577-2	Sequence 2, Appli
3	3763	100.0	3005	7	US-10-172-118-132	Sequence 132, App
4	3763	100.0	3005	8	US-10-342-887-132	Sequence 132, App
5	3763	100.0	3009	11	US-10-330-773-279	Sequence 279, App
6	3763	100.0	4293	5	US-09-720-086-3	Sequence 3, Appli
7	3763	100.0	4293	9	US-10-623-813-3	Sequence 3, Appli
8	3714	98.7	2318	9	US-10-623-813-83	Sequence 83, Appl
9	3714	98.7	4192	5	US-09-720-086-1	Sequence 1, Appli
10	3714	98.7	4192	9	US-10-623-813-1	Sequence 1, Appli
11	3714	98.7	6157	11	US-10-330-773-276	Sequence 276, App
12	2713	72.1	2938	7	US-10-264-237-953	Sequence 953, App
13	2169.5	57.7	4195	5	US-09-720-086-2	Sequence 2, Appli
14	2169.5	57.7	4195	9	US-10-623-813-2	Sequence 2, Appli
15	2157	57.3	4145	5	US-09-720-086-4	Sequence 4, Appli
16	2157	57.3	4145	6	US-10-144-577-3	Sequence 3, Appli
17	2157	57.3	4145	6	US-10-144-577-38	Sequence 38, Appl
18	2157	57.3	4145	7	US-10-172-118-1337	Sequence 1337, Ap
19	2157	57.3	4145	8	US-10-342-887-1337	Sequence 1337, Ap
20	2157	57.3	4145	9	US-10-623-813-4	Sequence 4, Appli
21	1878.5	49.9	3897	6	US-10-144-577-39	Sequence 39, Appl
22	1793.5	47.7	95484	11	US-10-330-773-275	Sequence 275, App
23	1771.5	47.1	127722	11	US-10-330-773-278	Sequence 278, App
24	1318	35.0	1060	6	US-10-106-698-926	Sequence 926, App
25	1095.5	29.1	709	3	US-09-969-034-1235	Sequence 1235, Ap
26	1003	26.7	711	10	US-10-450-763-21969	Sequence 21969, A
27	949	25.2	622	3	US-09-764-891-272	Sequence 272, App
c 28	942	25.0	994	12	US-10-301-480-606665	Sequence 606665,
c 29	942	25.0	994	12	US-10-301-480-1220074	Sequence 1220074,
30	872.5	23.2	2057	6	US-10-144-577-40	Sequence 40, Appl
31	785	20.9	2008	6	US-10-144-577-41	Sequence 41, Appl
32	750	19.9	440	5	US-09-720-086-54	Sequence 54, Appl
33	750	19.9	440	9	US-10-623-813-54	Sequence 54, Appl
34	693	18.4	2951	10	US-10-450-763-21970	Sequence 21970, A
35	689	18.3	1705	8	US-10-479-334-35	Sequence 35, Appl
36	645.5	17.2	1670	8	US-10-479-334-21	Sequence 21, Appl
37	640	17.0	1521	8	US-10-479-334-7	Sequence 7, Appli
38	636	16.9	522	5	US-09-720-086-61	Sequence 61, Appl
39	636	16.9	522	9	US-10-623-813-61	Sequence 61, Appl
40	629	16.7	598	6	US-10-027-632-199022	Sequence 199022,
41	629	16.7	598	6	US-10-027-632-199023	Sequence 199023,
42	629	16.7	598	7	US-10-027-632-199022	Sequence 199022,
43	629	16.7	598	7	US-10-027-632-199023	Sequence 199023,
44	609	16.2	435	5	US-09-720-086-18	Sequence 18, Appl
45	609	16.2	435	9	US-10-623-813-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-10-623-813-84
; Sequence 84, Application US/10623813
; Publication No. US20040234997A1
; GENERAL INFORMATION:
; APPLICANT: Li, En
; APPLICANT: Okano, Masaki
; APPLICANT: Xie, Shaoping
; APPLICANT: Chen, Taiping
; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0609.4560003
; CURRENT APPLICATION NUMBER: US/10/623,813
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 09/720,086
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/US99/14373
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,906
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 60/093,993
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 84

; LENGTH: 2371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-813-84

Alignment Scores:

Pred. No.:	0	Length:	2371
Score:	3763.00	Matches:	689
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-623-813-86 (1-689) x US-10-623-813-84 (1-2371)

```
Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 276

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 336

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 396

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 456

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 516

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 576

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 636

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 696

Qy        161 ProProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 756

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 816

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 876

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 936

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 996

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1056

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1057 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1116

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1176

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1177 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1236

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1237 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1296

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 1356

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1357 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 1416

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1417 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 1476

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 1536

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1537 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 1596

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 1656

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1657 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 1716

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1717 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 1776

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1777 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 1836

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1837 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 1896

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1897 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 1956

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1957 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2016

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2017 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2076

Qy        621 LysAspGlnHisProProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2077 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2136

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2137 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2196
```

```
Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2197 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2256

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              |||||||||||||||||||||||||||
Db       2257 CCGCTGAAGGAGTATTTGCGTGTGTG 2283

RESULT 2
US-10-144-577-2
; Sequence 2, Application US/10144577
; Publication No. US20030083292A1
; GENERAL INFORMATION:
;  APPLICANT: MacLeod, Alan Robert
;  TITLE OF INVENTION: Inhibitors of DNA Methyltransferase Isoforms
;  FILE REFERENCE: MET-005
;  CURRENT APPLICATION NUMBER: US/10/144,577
;  CURRENT FILING DATE:  2002-05-13
;  PRIOR APPLICATION NUMBER: US 60/290,202
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US 60/290,212
;  PRIOR FILING DATE: 2001-05-11
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 3005
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-144-577-2

Alignment Scores:
Pred. No.:                0          Length:          3005
Score:                    3763.00    Matches:         689
Percent Similarity:       100.0%     Conservative:    0
Best Local Similarity:    100.0%     Mismatches:      0
Query Match:              100.0%     Indels:          0
DB:                       6          Gaps:            0

US-10-623-813-86 (1-689) x US-10-144-577-2 (1-3005)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 958

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1018

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1078

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1138

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1198

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1258

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1318

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1378

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1438

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1439 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1498

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1558

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1559 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1618

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1619 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1678

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1679 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1738

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1739 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1798

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1799 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1858

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1859 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1918

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1919 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1978

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1979 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2038

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2039 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2098

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2099 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2158

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2159 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2218

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2278

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2279 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2338

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2339 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2398

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2399 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2458

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
```

```
Db       2459 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2518

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560

Db       2519 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2578

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580

Db       2579 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2638

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600

Db       2639 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2698

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620

Db       2699 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2758

Qy        621 LysAspGlnHisProProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640

Db       2759 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2818

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660

Db       2819 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2878

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680

Db       2879 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2938

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689

Db       2939 CCGCTGAAGGAGTATTTTGCGTGTGTG 2965

RESULT 3
US-10-172-118-132
; Sequence 132, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Chris
;  APPLICANT:  Van 't Veer, Laura
;  APPLICANT:  Van de Vijver, Marc
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-175-999
;  CURRENT APPLICATION NUMBER: US/10/172,118
;  CURRENT FILING DATE:  2002-06-14
;  PRIOR APPLICATION NUMBER: 60/380,770
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 132
;   LENGTH: 3005
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;  PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: AF067972
;   DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-132

Alignment Scores:
Pred. No.:              0           Length:         3005
Score:                  3763.00     Matches:        689
Percent Similarity:     100.0%      Conservative:   0
Best Local Similarity:  100.0%      Mismatches:     0
Query Match:            100.0%      Indels:         0
DB:                     7           Gaps:           0

US-10-623-813-86 (1-689) x US-10-172-118-132 (1-3005)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20

Db        899 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 958

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40

Db        959 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1018

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60

Db       1019 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1078

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80

Db       1079 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1138

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100

Db       1139 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1198

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120

Db       1199 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1258

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140

Db       1259 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1318

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160

Db       1319 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1378

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180

Db       1379 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1438

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200

Db       1439 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1498

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220

Db       1499 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1558

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240

Db       1559 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1618

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260

Db       1619 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1678

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280

Db       1679 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1738

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300

Db       1739 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1798

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320

Db       1799 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1858

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340

Db       1859 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1918

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360

Db       1919 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1978
```

```
Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1979 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2038

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2039 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2098

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2099 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2158

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2159 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2218

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2278

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2279 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2338

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2339 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2398

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2399 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2458

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2459 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2518

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2519 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2578

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2579 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2638

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2639 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2698

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2699 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2758

Qy        621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2759 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2818

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2819 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2878

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2879 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2938

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              ||||||||||||||||||||||||||||
Db       2939 CCGCTGAAGGAGTATTTTGCGTGTGTG 2965
```

```
RESULT 4
US-10-342-887-132
; Sequence 132, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter S.
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Christopher J.
;  APPLICANT:  Van 't Veer, Laura Johanna
;  APPLICANT:  Van de Vijver, Marc J.
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-188-999
;  CURRENT APPLICATION NUMBER: US/10/342,887
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: 60/298,918
;  PRIOR FILING DATE: 2001-06-18
;  PRIOR APPLICATION NUMBER: 60/380,710
;  PRIOR FILING DATE: 2002-05-14
;  PRIOR APPLICATION NUMBER: 10/172,118
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 132
;   LENGTH: 3005
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-132

Alignment Scores:
Pred. No.:                0              Length:          3005
Score:                    3763.00        Matches:         689
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       8              Gaps:            0

US-10-623-813-86 (1-689) x US-10-342-887-132 (1-3005)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 958

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1018

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1078

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1138

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1198

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1258

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1318

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1378

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1438

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1439 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1498

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220

Db       1499 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1558

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240

Db       1559 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1618

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260

Db       1619 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1678

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280

Db       1679 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1738

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300

Db       1739 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1798

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320

Db       1799 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1858

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340

Db       1859 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1918

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360

Db       1919 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1978

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380

Db       1979 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2038

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400

Db       2039 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2098

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420

Db       2099 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2158

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440

Db       2159 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2218

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460

Db       2219 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2278

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480

Db       2279 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2338

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500

Db       2339 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2398

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520

Db       2399 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2458

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540

Db       2459 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2518

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560

Db       2519 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2578

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580

Db       2579 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2638

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600

Db       2639 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2698

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620

Db       2699 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2758

Qy        621 LysAspGlnHisProProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640

Db       2759 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2818

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660

Db       2819 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2878

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680

Db       2879 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2938

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689

Db       2939 CCGCTGAAGGAGTATTTTGCGTGTGTG 2965

RESULT 5
US-10-330-773-279
; Sequence 279, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 279
;   LENGTH: 3009
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-330-773-279

Alignment Scores:
Pred. No.:                0              Length:          3009
Score:                    3763.00        Matches:         689
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       11             Gaps:            0

US-10-623-813-86 (1-689) x US-10-330-773-279 (1-3009)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20

Db        903 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 962

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40

Db        963 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1022

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60

Db       1023 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1082

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80

Db       1083 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1142
```

```
Qy        81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1143 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1202

Qy       101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1203 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1262

Qy       121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1263 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1322

Qy       141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1323 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1382

Qy       161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1383 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1442

Qy       181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1443 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1502

Qy       201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1503 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1562

Qy       221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1563 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1622

Qy       241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1623 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1682

Qy       261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1683 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1742

Qy       281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1743 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1802

Qy       301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1803 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1862

Qy       321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1863 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1922

Qy       341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1923 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1982

Qy       361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1983 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2042

Qy       381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2043 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2102

Qy       401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2103 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2162

Qy       421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2163 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2222

Qy       441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2223 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2282

Qy       461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2283 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2342

Qy       481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2343 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2402

Qy       501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2403 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2462

Qy       521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2463 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2522

Qy       541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2523 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2582

Qy       561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2583 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2642

Qy       581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2643 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2702

Qy       601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2703 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2762

Qy       621 LysAspGlnHisProProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2763 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2822

Qy       641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2823 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2882

Qy       661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2883 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2942

Qy       681 ProLeuLysGluTyrPheAlaCysVal 689
             ||||||||||||||||||||||||||
Db      2943 CCGCTGAAGGAGTATTTTGCGTGTGTG 2969
```

RESULT 6
US-09-720-086-3
; Sequence 3, Application US/09720086
; Publication No. US20060084053A1
; GENERAL INFORMATION:
; APPLICANT: Li, En
; APPLICANT: Okano, Masaki
; APPLICANT: Xie, Shaoping
; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,
; TITLE OF INVENTION: Polypeptides & Uses Thereof
; FILE REFERENCE: 0609.4560002
; CURRENT APPLICATION NUMBER: US/09/720,086
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/14373
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/090,906
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/093,993
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3
; LENGTH: 4293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-720-086-3

Alignment Scores:

Pred. No.:	0	Length:	4293
Score:	3763.00	Matches:	689
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-623-813-86 (1-689) x US-09-720-086-3 (1-4293)

```
Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 958

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1018

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1078

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1138

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1198

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1258

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1318

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1378

Qy        161 ProProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1438

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1439 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1498

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1558

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1559 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1618

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1619 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1678

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1679 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1738

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1739 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1798

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1799 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1858

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1859 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1918

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1919 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1978

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1979 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2038

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2039 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2098

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2099 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2158

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2159 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2218

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2278

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2279 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2338

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2339 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2398

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2399 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2458

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2459 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2518

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2519 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2578

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2579 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2638

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2639 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2698

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2699 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2758

Qy        621 LysAspGlnHisProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2759 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2818

Qy        641 MetGluArgValPheGlyProPheValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2819 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2878

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2879 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2938

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              |||||||||||||||||||||||||||
Db       2939 CCGCTGAAGGAGTATTTTGCGTGTGTG 2965

RESULT 7
US-10-623-813-3
; Sequence 3, Application US/10623813
; Publication No. US20040234997A1
; GENERAL INFORMATION:
;  APPLICANT: Li, En
;  APPLICANT:  Okano, Masaki
;  APPLICANT:  Xie, Shaoping
;  APPLICANT:  Chen, Taiping
;  TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 0609.4560003
;  CURRENT APPLICATION NUMBER: US/10/623,813
;  CURRENT FILING DATE:  2003-07-22
;  PRIOR APPLICATION NUMBER: US 09/720,086
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: PCT/US99/14373
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: US 60/090,906
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: US 60/093,993
;  PRIOR FILING DATE: 1998-07-24
;  NUMBER OF SEQ ID NOS: 119
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
;   LENGTH: 4293
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-623-813-3

Alignment Scores:
Pred. No.:              0              Length:         4293
Score:                  3763.00        Matches:        689
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     9              Gaps:           0

US-10-623-813-86 (1-689) x US-10-623-813-3 (1-4293)

Qy           1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         899 ATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGC 958

Qy          21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         959 CCTCCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAG 1018

Qy          41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1019 CCCGTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAG 1078

Qy          61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1079 TACGAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTC 1138

Qy          81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1139 TCCTGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAA 1198

Qy         101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1199 GGCACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAG 1258

Qy         121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1259 CTGATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCC 1318

Qy         141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1319 ATGTACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTG 1378

Qy         161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1379 TTCCCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAAC 1438

Qy         181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1439 AAGCCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAG 1498

Qy         201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1499 CCACCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCT 1558

Qy         221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1559 GAGGCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAG 1618

Qy         241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1619 AAGCCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTG 1678

Qy         261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1679 CGGCAGAAGTGCCGGAACATTGAGGACATCTGCATCTCCTGTGGGAGCCTCAATGTTACC 1738

Qy         281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1739 CTGGAACACCCCCTCTTCGTTGGAGGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAG 1798

Qy         301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1799 TGTGCGTACCAGTACGACGACGACGGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGC 1858

Qy         321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1859 CGTGAGGTGCTCATGTGCGGAAACAACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTG 1918

Qy         341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1919 GACCTCTTGGTGGGGCCGGGGGCTGCCCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1978

Qy         361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1979 TACATGTGCGGGCACAAGGGTACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCC 2038

Qy         381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2039 CGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTTAC 2098

Qy         401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2099 CCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATC 2158

Qy         421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2159 GCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGACCGCTACATTGCCTCG 2218

Qy         441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2219 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTAC 2278

Qy         461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2279 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTG 2338
```

```
Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2339 GTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTC 2398

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2399 TACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2458

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2459 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2518

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2519 AGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAA 2578

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2579 GAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2638

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2639 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2698

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2699 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGC 2758

Qy        621 LysAspGlnHisProProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2759 AAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAA 2818

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2819 ATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTG 2878

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2879 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2938

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              |||||||||||||||||||||||||||
Db       2939 CCGCTGAAGGAGTATTTTGCGTGTGTG 2965

RESULT 8
US-10-623-813-83
; Sequence 83, Application US/10623813
; Publication No. US20040234997A1
; GENERAL INFORMATION:
;  APPLICANT: Li, En
;  APPLICANT:  Okano, Masaki
;  APPLICANT:  Xie, Shaoping
;  APPLICANT:  Chen, Taiping
;  TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 0609.4560003
;  CURRENT APPLICATION NUMBER: US/10/623,813
;  CURRENT FILING DATE:  2003-07-22
;  PRIOR APPLICATION NUMBER: US 09/720,086
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: PCT/US99/14373
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: US 60/090,906
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: US 60/093,993
;  PRIOR FILING DATE: 1998-07-24
;  NUMBER OF SEQ ID NOS: 119
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83
;   LENGTH: 2318
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-623-813-83
```

```
Alignment Scores:
Pred. No.:                0                Length:          2318
Score:                    3714.00          Matches:         679
Percent Similarity:       98.8%            Conservative:    2
Best Local Similarity:    98.5%            Mismatches:      8
Query Match:              98.7%            Indels:          0
DB:                       9                Gaps:            0

US-10-623-813-86 (1-689) x US-10-623-813-83 (1-2318)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              ||||||||||||||||||||||||      ||||||||||||||||||||||||||||||
Db        148 ATGAATGCTGTGGAAGAGAACCAGGCCTCTGGAGAGTCTCAGAAGGTGGAGGAGGCCAGC 207

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 CCTCCTGCTGTGCAGCAGCCCACGGACCCTGCTTCTCCGACTGTGGCCACCACCCCTGAG 267

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              |||||||||  ||||||||||||||||||||||||||||||  ||||||||||||||||
Db        268 CCAGTAGGAGGGGATGCTGGGGACAAGAATGCTACCAAAGCAGCCGACGATGAGCCTGAG 327

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TATGAGGATGGCCGGGGCTTTGGCATTGGAGAGCTGGTGTGGGGGAAACTTCGGGGCTTC 387

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TCCTGGTGGCCAGGCCGAATTGTGTCTTGGTGGATGACAGGCCGGAGCCGAGCAGCTGAA 447

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 GGCACTCGCTGGGTCATGTGGTTCGGAGATGGCAAGTTCTCAGTGGTGTGTGTGGAGAAG 507

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 CTCATGCCGCTGAGCTCCTTCTGCAGTGCATTCCACCAGGCCACCTACAACAAGCAGCCC 567

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 ATGTACCGCAAAGCCATCTACGAAGTCCTCCAGGTGGCCAGCAGCCGTGCCGGGAAGCTG 627

Qy        161 ProProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||  ||||||||||||||||||||||:::  ||||||||||||||||||||||
Db        628 TTTCCAGCTTGCCATGACAGTGATGAAAGTGACAGTGGCAAGGCTGTGGAAGTGCAGAAC 687

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 AAGCAGATGATTGAATGGGCCCTCGGTGGCTTCCAGCCCTCGGGTCCTAAGGGCCTGGAG 747

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACCAGAAGAAGAGAAGAATCCTTACAAGGAAGTTTACACCGACATGTGGGTGGAGCCT 807

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||
Db        808 GAAGCAGCTGCTTACGCCCCACCCCCACCAGCCAAGAAACCCAGAAAGAGCACAACAGAG 867

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 AAACCTAAGGTCAAGGAGATCATTGATGAGCGCACAAGGGAGCGGCTGGTGTATGAGGTG 927

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 CGCCAGAAGTGCAGAAACATCGAGGACATTTGTATCTCATGTGGGAGCCTCAATGTCACC 987

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        988 CTGGAGCACCCACTCTTCATTGGAGGCATGTGCCAGAACTGTAAGAACTGCTTCTTGGAG 1047

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1048 TGTGCTTACCAGTATGACGACGATGGGTACCAGTCCTATTGCACCATCTGCTGTGGGGGG 1107

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 CGTGAAGTGCTCATGTGTGGGAACAACAACTGCTGCAGGTGCTTTTGTGTCGAGTGTGTG 1167

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 GATCTCTTGGTGGGGCCAGGAGCTGCTCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1227

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 TACATGTGCGGGCATAAGGGCACCTATGGGCTGCTGCGAAGACGGGAAGACTGGCCTTCT 1287

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 CGACTCCAGATGTTCTTTGCCAATAACCATGACCAGGAATTTGACCCCCCAAAGGTTTAC 1347

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CCACCTGTGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTCTTTGATGGGATT 1407

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTGGACCGCTACATTGCCTCC 1467

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGAAAGATCATGTAC 1527

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1528 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGACCTG 1587

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1588 GTGATTGGAGGCAGTCCCTGCAATGACCTCTCCATTGTCAACCCTGCCCGCAAGGGACTT 1647

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1648 TATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 1707

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1708 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 1767

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1768 AGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTGATGATTGACGCCAAA 1827

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1828 GAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAACAGG 1887

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1888 CCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCAGA 1947

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1948 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACTCTATAAAGCAGGGC 2007

Qy        621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2008 AAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATCCTGTGGTGCACTGAA 2067

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2068 ATGGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAACATGAGCCGCTTG 2127

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2128 GCGAGGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCT 2187

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              |||||||||||||||||||||||||||
Db       2188 CCGCTGAAGGAATATTTTGCTTGTGTG 2214

RESULT 9
US-09-720-086-1
; Sequence 1, Application US/09720086
; Publication No. US20060084053A1
; GENERAL INFORMATION:
;  APPLICANT: Li, En
;  APPLICANT:  Okano, Masaki
;  APPLICANT:  Xie, Shaoping
;  TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,
;  TITLE OF INVENTION:  Polypeptides & Uses Thereof
;  FILE REFERENCE: 0609.4560002
;  CURRENT APPLICATION NUMBER: US/09/720,086
;  CURRENT FILING DATE:  2000-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/14373
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: 60/090,906
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/093,993
;  PRIOR FILING DATE: 1998-07-24
;  NUMBER OF SEQ ID NOS: 82
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 4192
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: Unsure
;   LOCATION: (4161)..(4161)
;   OTHER INFORMATION: May be any nucleic acid
US-09-720-086-1

Alignment Scores:
Pred. No.:                0                Length:          4192
Score:                    3714.00          Matches:         679
Percent Similarity:       98.8%            Conservative:    2
Best Local Similarity:    98.5%            Mismatches:      8
Query Match:              98.7%            Indels:          0
DB:                       5                Gaps:            0

US-10-623-813-86 (1-689) x US-09-720-086-1 (1-4192)

Qy           1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
               ||||||||||||||||||||||||||      ||||||||||||||||||||||||||||||
Db         874 ATGAATGCTGTGGAAGAGAACCAGGCCTCTGGAGAGTCTCAGAAGGTGGAGGAGGCCAGC 933

Qy          21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         934 CCTCCTGCTGTGCAGCAGCCCACGGACCCTGCTTCTCCGACTGTGGCCACCACCCCTGAG 993

Qy          41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
               |||||||||  ||||||||||||||||||||||||||||||  ||||||||||||||||
Db         994 CCAGTAGGAGGGGATGCTGGGGACAAGAATGCTACCAAAGCAGCCGACGATGAGCCTGAG 1053

Qy          61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1054 TATGAGGATGGCCGGGGCTTTGGCATTGGAGAGCTGGTGTGGGGGAAACTTCGGGGCTTC 1113

Qy          81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1114 TCCTGGTGGCCAGGCCGAATTGTGTCTTGGTGGATGACAGGCCGGAGCCGAGCAGCTGAA 1173

Qy         101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1174 GGCACTCGCTGGGTCATGTGGTTCGGAGATGGCAAGTTCTCAGTGGTGTGTGTGGAGAAG 1233

Qy         121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1234 CTCATGCCGCTGAGCTCCTTCTGCAGTGCATTCCACCAGGCCACCTACAACAAGCAGCCC 1293
```

```
Qy       141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db      1294 ATGTACCGCAAAGCCATCTACGAAGTCCTCCAGGTGGCCAGCAGCCGTGCCGGGAAGCTG 1353

Qy       161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
Db      1354 TTTCCAGCTTGCCATGACAGTGATGAAAGTGACAGTGGCAAGGCTGTGGAAGTGCAGAAC 1413

Qy       181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db      1414 AAGCAGATGATTGAATGGGCCCTCGGTGGCTTCCAGCCCTCGGGTCCTAAGGGCCTGGAG 1473

Qy       201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db      1474 CCACCAGAAGAAGAGAAGAATCCTTACAAGGAAGTTTACACCGACATGTGGGTGGAGCCT 1533

Qy       221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
Db      1534 GAAGCAGCTGCTTACGCCCCACCCCCACCAGCCAAGAAACCCAGAAAGAGCACAACAGAG 1593

Qy       241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db      1594 AAACCTAAGGTCAAGGAGATCATTGATGAGCGCACAAGGGAGCGGCTGGTGTATGAGGTG 1653

Qy       261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db      1654 CGCCAGAAGTGCAGAAACATCGAGGACATTTGTATCTCATGTGGGAGCCTCAATGTCACC 1713

Qy       281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db      1714 CTGGAGCACCCACTCTTCATTGGAGGCATGTGCCAGAACTGTAAGAACTGCTTCTTGGAG 1773

Qy       301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1774 TGTGCTTACCAGTATGACGACGATGGGTACCAGTCCTATTGCACCATCTGCTGTGGGGGG 1833

Qy       321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db      1834 CGTGAAGTGCTCATGTGTGGGAACAACAACTGCTGCAGGTGCTTTTGTGTCGAGTGTGTG 1893

Qy       341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
Db      1894 GATCTCTTGGTGGGGCCAGGAGCTGCTCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1953

Qy       361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
Db      1954 TACATGTGCGGGCATAAGGGCACCTATGGGCTGCTGCGAAGACGGGAAGACTGGCCTTCT 2013

Qy       381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db      2014 CGACTCCAGATGTTCTTTGCCAATAACCATGACCAGGAATTTGACCCCCCAAAGGTTTAC 2073

Qy       401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      2074 CCACCTGTGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTCTTTGATGGGATT 2133

Qy       421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      2134 GCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTGGACCGCTACATTGCCTCC 2193

Qy       441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      2194 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGAAAGATCATGTAC 2253

Qy       461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db      2254 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGACCTG 2313

Qy       481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db      2314 GTGATTGGAGGCAGTCCCTGCAATGACCTCTCCATTGTCAACCCTGCCCGCAAGGGACTT 2373

Qy       501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      2374 TATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2433

Qy       521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
Db      2434 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2493

Qy       541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db      2494 AGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTGATGATTGACGCCAAA 2553

Qy       561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db      2554 GAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAACAGG 2613

Qy       581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
Db      2614 CCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCAGA 2673

Qy       601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db      2674 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACTCTATAAAGCAGGGC 2733

Qy       621 LysAspGlnHisProProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db      2734 AAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATCCTGTGGTGCACTGAA 2793

Qy       641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2794 ATGGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAACATGAGCCGCTTG 2853

Qy       661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db      2854 GCGAGGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCT 2913

Qy       681 ProLeuLysGluTyrPheAlaCysVal 689
Db      2914 CCGCTGAAGGAATATTTTGCTTGTGTG 2940
```

RESULT 10
US-10-623-813-1
; Sequence 1, Application US/10623813
; Publication No. US20040234997A1
; GENERAL INFORMATION:
; APPLICANT: Li, En
; APPLICANT: Okano, Masaki
; APPLICANT: Xie, Shaoping
; APPLICANT: Chen, Taiping
; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0609.4560003
; CURRENT APPLICATION NUMBER: US/10/623,813
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 09/720,086
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/US99/14373
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,906
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 60/093,993
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 4192
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4161)..(4161)
; OTHER INFORMATION: n is a, c, g, or t
US-10-623-813-1

Alignment Scores:

Pred. No.:	0	Length:	4192
Score:	3714.00	Matches:	679
Percent Similarity:	98.8%	Conservative:	2
Best Local Similarity:	98.5%	Mismatches:	8
Query Match:	98.7%	Indels:	0
DB:	9	Gaps:	0

US-10-623-813-86 (1-689) x US-10-623-813-1 (1-4192)

```
Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              |||||||||||||||||||||||||      |||||||||||||||||||||||||||||
Db        874 ATGAATGCTGTGGAAGAGAACCAGGCCTCTGGAGAGTCTCAGAAGGTGGAGGAGGCCAGC 933

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        934 CCTCCTGCTGTGCAGCAGCCCACGGACCCTGCTTCTCCGACTGTGGCCACCACCCCTGAG 993

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              |||||||||    |||||||||||||||||||||||||||||    ||||||||||||||||
Db        994 CCAGTAGGAGGGGATGCTGGGGACAAGAATGCTACCAAAGCAGCCGACGATGAGCCTGAG 1053

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1054 TATGAGGATGGCCGGGGCTTTGGCATTGGAGAGCTGGTGTGGGGGAAACTTCGGGGCTTC 1113

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1114 TCCTGGTGGCCAGGCCGAATTGTGTCTTGGTGGATGACAGGCCGGAGCCGAGCAGCTGAA 1173

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1174 GGCACTCGCTGGGTCATGTGGTTCGGAGATGGCAAGTTCTCAGTGGTGTGTGTGGAGAAG 1233

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1234 CTCATGCCGCTGAGCTCCTTCTGCAGTGCATTCCACCAGGCCACCTACAACAAGCAGCCC 1293

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1294 ATGTACCGCAAAGCCATCTACGAAGTCCTCCAGGTGGCCAGCAGCCGTGCCGGGAAGCTG 1353

Qy        161 ProProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||    |||||||||||||||||||||:::    ||||||||||||||||||||
Db       1354 TTTCCAGCTTGCCATGACAGTGATGAAAGTGACAGTGGCAAGGCTGTGGAAGTGCAGAAC 1413

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1414 AAGCAGATGATTGAATGGGCCCTCGGTGGCTTCCAGCCCTCGGGTCCTAAGGGCCTGGAG 1473

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1474 CCACCAGAAGAAGAGAAGAATCCTTACAAGGAAGTTTACACCGACATGTGGGTGGAGCCT 1533

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    |||
Db       1534 GAAGCAGCTGCTTACGCCCCACCCCCACCAGCCAAGAAACCCAGAAAGAGCACAACAGAG 1593

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1594 AAACCTAAGGTCAAGGAGATCATTGATGAGCGCACAAGGGAGCGGCTGGTGTATGAGGTG 1653

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1654 CGCCAGAAGTGCAGAAACATCGAGGACATTTGTATCTCATGTGGGAGCCTCAATGTCACC 1713

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db       1714 CTGGAGCACCCACTCTTCATTGGAGGCATGTGCCAGAACTGTAAGAACTGCTTCTTGGAG 1773

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1774 TGTGCTTACCAGTATGACGACGATGGGTACCAGTCCTATTGCACCATCTGCTGTGGGGGG 1833

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1834 CGTGAAGTGCTCATGTGTGGGAACAACAACTGCTGCAGGTGCTTTTGTGTCGAGTGTGTG 1893

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1894 GATCTCTTGGTGGGGCCAGGAGCTGCTCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1953

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1954 TACATGTGCGGGCATAAGGGCACCTATGGGCTGCTGCGAAGACGGGAAGACTGGCCTTCT 2013

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2014 CGACTCCAGATGTTCTTTGCCAATAACCATGACCAGGAATTTGACCCCCCAAAGGTTTAC 2073

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2074 CCACCTGTGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTCTTTGATGGGATT 2133

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2134 GCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTGGACCGCTACATTGCCTCC 2193

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2194 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGAAAGATCATGTAC 2253

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2254 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGACCTG 2313

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2314 GTGATTGGAGGCAGTCCCTGCAATGACCTCTCCATTGTCAACCCTGCCCGCAAGGGACTT 2373

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2374 TATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2433

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2434 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2493

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2494 AGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTGATGATTGACGCCAAA 2553

Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2554 GAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAACAGG 2613

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2614 CCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCAGA 2673

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2674 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACTCTATAAAGCAGGGC 2733

Qy        621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2734 AAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATCCTGTGGTGCACTGAA 2793

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2794 ATGGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAACATGAGCCGCTTG 2853

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2854 GCGAGGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCT 2913
```

```
Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              |||||||||||||||||||||||||||
Db       2914 CCGCTGAAGGAATATTTTGCTTGTGTG 2940

RESULT 11
US-10-330-773-276
; Sequence 276, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 276
;   LENGTH: 6157
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-330-773-276

Alignment Scores:
Pred. No.:              0           Length:         6157
Score:                  3714.00     Matches:        679
Percent Similarity:     98.8%       Conservative:   2
Best Local Similarity:  98.5%       Mismatches:     8
Query Match:            98.7%       Indels:         0
DB:                     11          Gaps:           0

US-10-623-813-86 (1-689) x US-10-330-773-276 (1-6157)

Qy          1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
              |||||||||||||||||||||||||      ||||||||||||||||||||||||||||||
Db        866 ATGAATGCTGTGGAAGAGAACCAGGCCTCTGGAGAGTCTCAGAAGGTGGAGGAGGCCAGC 925

Qy         21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 CCTCCTGCTGTGCAGCAGCCCACGGACCCTGCTTCTCCGACTGTGGCCACCACCCCTGAG 985

Qy         41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
              |||||||||   ||||||||||||||||||||||||||||||   ||||||||||||||||
Db        986 CCAGTAGGAGGGGATGCTGGGGACAAGAATGCTACCAAAGCAGCCGACGATGAGCCTGAG 1045

Qy         61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 TATGAGGATGGCCGGGGCTTTGGCATTGGAGAGCTGGTGTGGGGGAAACTTCGGGGCTTC 1105

Qy         81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1106 TCCTGGTGGCCAGGCCGAATTGTGTCTTGGTGGATGACAGGCCGGAGCCGAGCAGCTGAA 1165

Qy        101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 GGCACTCGCTGGGTCATGTGGTTCGGAGATGGCAAGTTCTCAGTGGTGTGTGTGGAGAAG 1225

Qy        121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 CTCATGCCGCTGAGCTCCTTCTGCAGTGCATTCCACCAGGCCACCTACAACAAGCAGCCC 1285

Qy        141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 ATGTACCGCAAAGCCATCTACGAAGTCCTCCAGGTGGCCAGCAGCCGTGCCGGGAAGCTG 1345

Qy        161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
              ||||||   |||||||||||||||||||||||:::   ||||||||||||||||||||||
Db       1346 TTTCCAGCTTGCCATGACAGTGATGAAAGTGACAGTGGCAAGGCTGTGGAAGTGCAGAAC 1405

Qy        181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1406 AAGCAGATGATTGAATGGGCCCTCGGTGGCTTCCAGCCCTCGGGTCCTAAGGGCCTGGAG 1465

Qy        201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1466 CCACCAGAAGAAGAGAAGAATCCTTACAAGGAAGTTTACACCGACATGTGGGTGGAGCCT 1525

Qy        221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db       1526 GAAGCAGCTGCTTACGCCCCACCCCCACCAGCCAAGAAACCCAGAAAGAGCACAACAGAG 1585

Qy        241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1586 AAACCTAAGGTCAAGGAGATCATTGATGAGCGCACAAGGGAGCGGCTGGTGTATGAGGTG 1645

Qy        261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1646 CGCCAGAAGTGCAGAAACATCGAGGACATTTGTATCTCATGTGGGAGCCTCAATGTCACC 1705

Qy        281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db       1706 CTGGAGCACCCACTCTTCATTGGAGGCATGTGCCAGAACTGTAAGAACTGCTTCTTGGAG 1765

Qy        301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1766 TGTGCTTACCAGTATGACGACGATGGGTACCAGTCCTATTGCACCATCTGCTGTGGGGGG 1825

Qy        321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1826 CGTGAAGTGCTCATGTGTGGGAACAACAACTGCTGCAGGTGCTTTTGTGTCGAGTGTGTG 1885

Qy        341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1886 GATCTCTTGGTGGGGCCAGGAGCTGCTCAGGCAGCCATTAAGGAAGACCCCTGGAACTGC 1945

Qy        361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1946 TACATGTGCGGGCATAAGGGCACCTATGGGCTGCTGCGAAGACGGGAAGACTGGCCTTCT 2005

Qy        381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2006 CGACTCCAGATGTTCTTTGCCAATAACCATGACCAGGAATTTGACCCCCCAAAGGTTTAC 2065

Qy        401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db.      2066 CCACCTGTGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTCTTTGATGGGATT 2125

Qy        421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2126 GCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTGGACCGCTACATTGCCTCC 2185

Qy        441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2186 GAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGGAAAGATCATGTAC 2245

Qy        461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2246 GTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGACCTG 2305

Qy        481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2306 GTGATTGGAGGCAGTCCCTGCAATGACCTCTCCATTGTCAACCCTGCCCGCAAGGGACTT 2365

Qy        501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2366 TATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCC 2425

Qy        521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2426 AAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTT 2485

Qy        541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2486 AGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTGATGATTGACGCCAAA 2545
```

```
Qy        561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2546 GAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAACAGG 2605

Qy        581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2606 CCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCAGA 2665

Qy        601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2666 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACTCTATAAAGCAGGGC 2725

Qy        621 LysAspGlnHisProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2726 AAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATCCTGTGGTGCACTGAA 2785

Qy        641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2786 ATGGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAACATGAGCCGCTTG 2845

Qy        661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2846 GCGAGGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCT 2905

Qy        681 ProLeuLysGluTyrPheAlaCysVal 689
              |||||||||||||||||||||||||||
Db       2906 CCGCTGAAGGAATATTTTGCTTGTGTG 2932

RESULT 12
US-10-264-237-953
; Sequence 953, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA131P1
;  CURRENT APPLICATION NUMBER: US/10/264,237
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/16450
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: US 60/205,515
;  PRIOR FILING DATE: 2000-05-19
;  NUMBER OF SEQ ID NOS: 2876
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 953
;   LENGTH: 2938
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (491)..(491)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1987)..(1987)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1994)..(1994)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2892)..(2892)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2932)..(2932)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-953

Alignment Scores:
Pred. No.:              2.11e-293       Length:         2938
Score:                  2713.00         Matches:        496
Percent Similarity:     99.0%           Conservative:   0
Best Local Similarity:  99.0%           Mismatches:     5
Query Match:            72.1%           Indels:         0
DB:                     7               Gaps:           0

US-10-623-813-86 (1-689) x US-10-264-237-953 (1-2938)

Qy        189 GlyGlyPheGlnProSerGlyProLysGlyLeuGluProProGluGluGluLysAsnPro 208
              |||     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGCCAGCCTTCTGGCCCTAAGGGCCTAGAGCCACCAGAAGAAGAGAAGAATCCC 60

Qy        209 TyrLysGluValTyrThrAspMetTrpValGluProGluAlaAlaAlaTyrAlaProPro 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TACAAAGAAGTGTACACGGACATGTGGGTGGAACCTGAGGCAGCTGCCTACGCACCACCT 120

Qy        229 ProProAlaLysLysProArgLysSerThrAlaGluLysProLysValLysGluIleIle 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAGAAGCCCAAGGTCAAGGAGATTATT 180

Qy        249 AspGluArgThrArgGluArgLeuValTyrGluValArgGlnLysCysArgAsnIleGlu 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTGCGGCAGAAGTGCCGGAACATTGAG 240

Qy        269 AspIleCysIleSerCysGlySerLeuAsnValThrLeuGluHisProLeuPheValGly 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GACATCTGCATCTCCTGTGGGAGCCTCAATGTTACCCTGGAACACCCCCTCTTCGTTGGA 300

Qy        289 GlyMetCysGlnAsnCysLysAsnCysPheLeuGluCysAlaTyrGlnTyrAspAspAsp 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAATGTGCCAAAACTGCAAGAACTGCTTTCTGGAGTGTGCGTACCAGTACGACGACGAC 360

Qy        309 GlyTyrGlnSerTyrCysThrIleCysCysGlyGlyArgGluValLeuMetCysGlyAsn 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCTACCAGTCCTACTGCACCATCTGCTGTGGGGGCCGTGAGGTGCTCATGTGCGGAAAC 420

Qy        329 AsnAsnCysCysArgCysPheCysValGluCysValAspLeuLeuValGlyProGlyAla 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACAACTGCTGCAGGTGCTTTTGCGTGGAGTGTGTGGACCTCTTGGTGGGGCCGGGGGCT 480

Qy        349 AlaGlnAlaAlaIleLysGluAspProTrpAsnCysTyrMetCysGlyHisLysGlyThr 368
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GSCCARGCAGNCATTAAGGAAGACCCCTGGAACTGCTACATGTGCGGGCACAAGGGTACC 540

Qy        369 TyrGlyLeuLeuArgArgArgGluAspTrpProSerArgLeuGlnMetPhePheAlaAsn 388
              ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db        541 TACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCCCGGYTCCAGATGTTCTTCGCTAAT 600

Qy        389 AsnHisAspGlnGluPheAspProProLysValTyrProProValProAlaGluLysArg 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACCACGACCAGGAATTTGACCCTCCAAAGGTTTACCCACCTGTCCCAGCTGAGAAGAGG 660

Qy        409 LysProIleArgValLeuSerLeuPheAspGlyIleAlaThrGlyLeuLeuValLeuLys 428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATCGCTACAGGGCTCCTGGTGCTGAAG 720

Qy        429 AspLeuGlyIleGlnValAspArgTyrIleAlaSerGluValCysGluAspSerIleThr 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTTGGGCATTCAGGTGGACCGCTACATTGCCTCGGARGTGTGTGAGGACTCCATCACG 780

Qy        449 ValGlyMetValArgHisGlnGlyLysIleMetTyrValGlyAspValArgSerValThr 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGGGCATGGTGCGGCACCAGGGGAAGATCATGTACGTCGGGGACGTCCGCAGCGTCACA 840

Qy        469 GlnLysHisIleGlnGluTrpGlyProPheAspLeuValIleGlyGlySerProCysAsn 488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTGGTGATTGGGGGCAGTCCCTGCAAT 900

Qy        489 AspLeuSerIleValAsnProAlaArgLysGlyLeuTyrGluGlyThrGlyArgLeuPhe 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACCTCTCCATCGTCAACCCTGCTCGCAAGGGCCTCTACGAGGGCACTGGCCGGCTCTTC 960

Qy        509 PheGluPheTyrArgLeuLeuHisAspAlaArgProLysGluGlyAspAspArgProPhe 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        961 TTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCCAAGGAGGGAGATGATCGCCCCTTC 1020

Qy        529 PheTrpLeuPheGluAsnValValAlaMetGlyValSerAspLysArgAspIleSerArg 548

Db       1021 TTCTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTTAGTGACAAGAGGGACATCTCGCGA 1080

Qy        549 PheLeuGluSerAsnProValMetIleAspAlaLysGluValSerAlaAlaHisArgAla 568

Db       1081 TTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAAGAAGTGTCAGCTGCACACAGGGCC 1140

Qy        569 ArgTyrPheTrpGlyAsnLeuProGlyMetAsnArgProLeuAlaSerThrValAsnAsp 588

Db       1141 CGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGGCCGTTGGCATCCACTGTGAATGAT 1200

Qy        589 LysLeuGluLeuGlnGluCysLeuGluHisGlyArgIleAlaLysPheSerLysValArg 608

Db       1201 AAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGGATAGCCAAGTTCAGCAAAGTGAGG 1260

Qy        609 ThrIleThrThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisProValPhe 628

Db       1261 ACCATTACTACGAGGTCAAACTCCATAAAGCAGGGCAAAGACCAGCATTTTCCTGTCTTC 1320

Qy        629 MetAsnGluLysGluAspIleLeuTrpCysThrGluMetGluArgValPheGlyPhePro 648

Db       1321 ATGAATGAGAAAGAGGACATCTTATGGTGCACTGAAATGGAAAGGGTATTTGGTTTCCCA 1380

Qy        649 ValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuLeuGlyArg 668

Db       1381 GTCCACTATACTGACGTCTCCAACATGAGCCGCTTGGCGAGGCAGAGACTGCTGGGCCGG 1440

Qy        669 SerTrpSerValProValIleArgHisLeuPheAlaProLeuLysGluTyrPheAlaCys 688

Db       1441 TCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCTCCGCTGAAGGAGTATTTTGCGTGT 1500

Qy        689 Val 689

Db       1501 GTG 1503

RESULT 13
US-09-720-086-2
; Sequence 2, Application US/09720086
; Publication No. US20060084053A1
; GENERAL INFORMATION:
;  APPLICANT: Li, En
;  APPLICANT:  Okano, Masaki
;  APPLICANT:  Xie, Shaoping
;  TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,
;  TITLE OF INVENTION:  Polypeptides & Uses Thereof
;  FILE REFERENCE: 0609.4560002
;  CURRENT APPLICATION NUMBER: US/09/720,086
;  CURRENT FILING DATE:  2000-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/14373
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: 60/090,906
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/093,993
;  PRIOR FILING DATE: 1998-07-24
;  NUMBER OF SEQ ID NOS: 82
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 4195
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-720-086-2

Alignment Scores:
Pred. No.:                4.92e-232       Length:          4195
Score:                    2169.50         Matches:         404
Percent Similarity:       71.9%           Conservative:    100
Best Local Similarity:    57.6%           Mismatches:      156
Query Match:              57.7%           Indels:          41
DB:                       5               Gaps:            6
```

```
US-10-623-813-86 (1-689) x US-09-720-086-2 (1-4195)

Qy         10 ProGlyGluSerGlnLysValGluGluAlaSerProProAlaValGlnGlnProThrAsp 29

Db        797 CCTGCCAGCGTCGACTTCATGGAAGAAGTGACACCTAAGAGCGTCAGTACCCCA------ 850

Qy         30 ProAlaSerProThrValAlaThrThrProGluProValGlySerAspAlaGlyAspLys 49

Db        851 ------TCAGTTGACTTGAGCCAGGATGGAGATCAGGAGGGTATGGATACCACACAGGTG 904

Qy         50 AsnAlaThrLysAlaGlyAspAspGluProGluTyrGluAspGlyArgGlyPheGlyIle 69

Db        905 GATGCAGAGAGCAGAGATGGAGACAGCACAGAGTATCAGGATGATAAAGAGTTTGGAATA 964

Qy         70 GlyGluLeuValTrpGlyLysLeuArgGlyPheSerTrpTrpProGlyArgIleValSer 89

Db        965 GGTGACCTCGTGTGGGGAAAGATCAAGGGCTTCTCCTGGTGGCCTGCCATGGTGGTGTCC 1024

Qy         90 TrpTrpMetThrGlyArgSerArgAlaAlaGluGlyThrArgTrpValMetTrpPheGly 109

Db       1025 TGGAAAGCCACCTCCAAGCGACAGGCCATGCCCGGAATGCGCTGGGTACAGTGGTTTGGT 1084

Qy        110 AspGlyLysPheSerValValCysValGluLysLeuMetProLeuSerSerPheCysSer 129

Db       1085 GATGGCAAGTTTTCTGAGATCTCTGCTGACAAACTGGTGGCTCTGGGGCTGTTCAGCCAG 1144

Qy        130 AlaPheHisGlnAlaThrTyrAsnLysGlnProMetTyrArgLysAlaIleTyrGluVal 149

Db       1145 CACTTTAATCTGGCTACCTTCAATAAGCTGGTTTCTTATAGGAAGGCCATGTACCACACT 1204

Qy        150 LeuGlnValAlaSerSerArgAlaGlyLysLeuProValCysHisAspSerAspGlu 169

Db       1205 CTGGAGAAAGCCAGGGTTCGAGCTGGCAAGACCTTC--------------------TCC 1243

Qy        170 SerAspThrAlaLysAlaValGluValGlnAsnLysProMetIleGluTrpAlaLeuGly 189

Db       1244 AGCAGTCCTGGAGAGTCACTGGAGGACCAGCTGAAGCCCATGCTGGAGTGGGCCCACGGT 1303

Qy        190 GlyPheGlnProSerGlyProLysGlyLeuGluPro---------------------- 201

Db       1304 GGCTTCAAGCCTACTGGGATCGAGGGCCTCAAACCCAACAAGAAGCAACCAGTGGTTAAT 1363

Qy        202 ------------------------------------ProGluGluGluLysAsnPro 208

Db       1364 AAGTCGAAGGTGCGTCGTTCAGACAGTAGGAACTTAGAACCCAGGAGACGCGAGAACAAA 1423

Qy        209 TyrLysGluValTyrThrAspMetTrpValGluProGluAlaAlaAlaTyrAlaProPro 228

Db       1424 AGTCGAAGACGCACAACCAAT--------------GACTCTGCTGCTTCTGAGTCCCCC 1468

Qy        229 ProProAlaLysLysProArgLysSerThrAlaGluLysProLysValLysGluIleIle 248

Db       1469 CCACCCAAGCGCCTCAAGACAAATAGCTATGGCGGGAAGGACCGAGGGGAG--------- 1519

Qy        249 AspGluArgThrArgGluArgLeuValTyrGluValArgGlnLysCysArgAsnIleGlu 268

Db       1520 GATGAGGAGAGCCGAGAACGGATGGCTTCTGAAGTCACCAACAACAAGGGCAATCTGGAA 1579

Qy        269 AspIleCysIleSerCysGlySerLeuAsnValThrLeuGluHisProLeuPheValGly 288

Db       1580 GACCGCTGTTTGTCCTGTGGAAAGAAGAACCCTGTGTCCTTCCACCCCCTCTTTGAGGGT 1639

Qy        289 GlyMetCysGlnAsnCysLysAsnCysPheLeuGluCysAlaTyrGlnTyrAspAspAsp 308

Db       1640 GGGCTCTGTCAGAGTTGCCGGGATCGCTTCCTAGAGCTCTTCTACATGTATGATGAGGAC 1699

Qy        309 GlyTyrGlnSerTyrCysThrIleCysCysGlyGlyArgGluValLeuMetCysGlyAsn 328

Db       1700 GGCTATCAGTCCTACTGCACCGTGTGCTGTGAGGGCCGTGAACTGCTGCTGTGCAGTAAC 1759

Qy        329 AsnAsnCysCysArgCysPheCysValGluCysValAspLeuLeuValGlyProGlyAla 348

Db       1760 ACAAGCTGCTGCAGATGCTTCTGTGTGGAGTGTCTGGAGGTGCTGGTGGGCGCAGGCACA 1819
```

```
Qy        349 AlaGlnAlaAlaIleLysGluAspProTrpAsnCysTyrMetCysGlyHisLysGlyThr 368
Db       1820 GCTGAGGATGCCAAGCTGCAGGAACCCTGGAGCTGCTATATGTGCCTCCCTCAGCGCTGC 1879

Qy        369 TyrGlyLeuLeuArgArgArgGluAspTrpProSerArgLeuGlnMetPhePheAlaAsn 388
Db       1880 CATGGGGTCCTCCGACGCAGGAAAGATTGGAACATGCGCCTGCAAGACTTCTTCACTACT 1939

Qy        389 AsnHisAsp---GlnGluPheAspProProLysValTyrProProValProAlaGluLys 407
Db       1940 GATCCTGACCTGGAAGAATTTGAGCCACCCAAGTTGTACCCAGCAATTCCTGCAGCCAAA 1999

Qy        408 ArgLysProIleArgValLeuSerLeuPheAspGlyIleAlaThrGlyLeuLeuValLeu 427
Db       2000 AGGAGGCCCATTAGAGTCCTGTCTCTGTTTGATGGAATTGCAACGGGGTACTTGGTGCTC 2059

Qy        428 LysAspLeuGlyIleGlnValAspArgTyrIleAlaSerGluValCysGluAspSerIle 447
Db       2060 AAGGAGTTGGGTATTAAAGTGGAAAAGTACATTGCCTCCGAAGTCTGTGCAGAGTCCATC 2119

Qy        448 ThrValGlyMetValArgHisGlnGlyLysIleMetTyrValGlyAspValArgSerVal 467
Db       2120 GCTGTGGGAACTGTTAAGCATGAAGGCCAGATCAAATATGTCAATGACGTCCGGAAAATC 2179

Qy        468 ThrGlnLysHisIleGlnGluTrpGlyProPheAspLeuValIleGlyGlySerProCys 487
Db       2180 ACCAAGAAAAATATTGAAGAGTGGGGCCCGTTCGACTTGGTGATTGGTGGAAGCCCATGC 2239

Qy        488 AsnAspLeuSerIleValAsnProAlaArgLysGlyLeuTyrGluGlyThrGlyArgLeu 507
Db       2240 AATGATCTCTCTAACGTCAATCCTGCCCGCAAAGGTTTATATGAGGGCACAGGAAGGCTC 2299

Qy        508 PhePheGluPheTyrArgLeuLeuHisAspAlaArgProLysGluGlyAspAspArgPro 527
Db       2300 TTCTTCGAGTTTTACCACTTGCTGAATTATACCCGCCCCAAGGAGGGCGACAACCGTCCA 2359

Qy        528 PhePheTrpLeuPheGluAsnValValAlaMetGlyValSerAspLysArgAspIleSer 547
Db       2360 TTCTTCTGGATGTTCGAGAATGTTGTGGCCATGAAAGTGAATGACAAGAAAGACATCTCA 2419

Qy        548 ArgPheLeuGluSerAsnProValMetIleAspAlaLysGluValSerAlaAlaHisArg 567
Db       2420 AGATTCCTGGCATGTAACCCAGTGATGATCGATGCCATCAAGGTGTCTGCTGCTCACAGG 2479

Qy        568 AlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArgProLeuAlaSerThrValAsn 587
Db       2480 GCCCGGTACTTCTGGGGTAACCTACCCGGAATGAACAGGCCCGTGATGGCTTCAAAGAAT 2539

Qy        588 AspLysLeuGluLeuGlnGluCysLeuGluHisGlyArgIleAlaLysPheSerLysVal 607
Db       2540 GATAAGCTCGAGCTGCAGGACTGCCTGGAGTTCAGTAGGACAGCAAAGTTAAAGAAAGTG 2599

Qy        608 ArgThrIleThrThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisPheProVal 627
Db       2600 CAGACAATAACCACCAAGTCGAACTCCATCAGACAGGGCAAAAACCAGCTTTTCCCTGTA 2659

Qy        628 PheMetAsnGluLysGluAspIleLeuTrpCysThrGluMetGluArgValPheGlyPhe 647
Db       2660 GTCATGAATGGCAAGGACGACGTTTTGTGGTGCACTGAGCTCGAAAGGATCTTCGGCTTC 2719

Qy        648 ProValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuLeuGly 667
Db       2720 CCTGCTCACTACACGGACGTGTCCAACATGGGCCGCGGCGCCCGTCAGAAGCTGCTGGGC 2779

Qy        668 ArgSerTrpSerValProValIleArgHisLeuPheAlaProLeuLysGluTyrPheAla 687
Db       2780 AGGTCCTGGAGTGTACCGGTCATCAGACACCTGTTTGCCCCCTTGAAGGACTACTTTGCC 2839

Qy        688 Cys 688
Db       2840 TGT 2842
```

```
RESULT 14
US-10-623-813-2
; Sequence 2, Application US/10623813
; Publication No. US20040234997A1
; GENERAL INFORMATION:
;  APPLICANT: Li, En
;  APPLICANT:  Okano, Masaki
;  APPLICANT:  Xie, Shaoping
;  APPLICANT:  Chen, Taiping
;  TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 0609.4560003
;  CURRENT APPLICATION NUMBER: US/10/623,813
;  CURRENT FILING DATE:  2003-07-22
;  PRIOR APPLICATION NUMBER: US 09/720,086
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: PCT/US99/14373
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: US 60/090,906
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: US 60/093,993
;  PRIOR FILING DATE: 1998-07-24
;  NUMBER OF SEQ ID NOS: 119
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;   LENGTH: 4195
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-623-813-2

Alignment Scores:
Pred. No.:                4.92e-232       Length:          4195
Score:                    2169.50         Matches:         404
Percent Similarity:       71.9%           Conservative:    100
Best Local Similarity:    57.6%           Mismatches:      156
Query Match:              57.7%           Indels:          41
DB:                       9               Gaps:            6

US-10-623-813-86 (1-689) x US-10-623-813-2 (1-4195)

Qy         10 ProGlyGluSerGlnLysValGluGluAlaSerProProAlaValGlnGlnProThrAsp 29
Db        797 CCTGCCAGCGTCGACTTCATGGAAGAAGTGACACCTAAGAGCGTCAGTACCCCA------ 850

Qy         30 ProAlaSerProThrValAlaThrThrProGluProValGlySerAspAlaGlyAspLys 49
Db        851 ------TCAGTTGACTTGAGCCAGGATGGAGATCAGGAGGGTATGGATACCACACAGGTG 904

Qy         50 AsnAlaThrLysAlaGlyAspAspGluProGluTyrGluAspGlyArgGlyPheGlyIle 69
Db        905 GATGCAGAGAGCAGAGATGGAGACAGCACAGAGTATCAGGATGATAAAGAGTTTGGAATA 964

Qy         70 GlyGluLeuValTrpGlyLysLeuArgGlyPheSerTrpTrpProGlyArgIleValSer 89
Db        965 GGTGACCTCGTGTGGGGAAAGATCAAGGGCTTCTCCTGGTGGCCTGCCATGGTGGTGTCC 1024

Qy         90 TrpTrpMetThrGlyArgSerArgAlaAlaGluGlyThrArgTrpValMetTrpPheGly 109
Db       1025 TGGAAAGCCACCTCCAAGCGACAGGCCATGCCCGGAATGCGCTGGGTACAGTGGTTTGGT 1084

Qy        110 AspGlyLysPheSerValValCysValGluLysLeuMetProLeuSerSerPheCysSer 129
Db       1085 GATGGCAAGTTTTCTGAGATCTCTGCTGACAAACTGGTGGCTCTGGGGCTGTTCAGCCAG 1144

Qy        130 AlaPheHisGlnAlaThrTyrAsnLysGlnProMetTyrArgLysAlaIleTyrGluVal 149
Db       1145 CACTTTAATCTGGCTACCTTCAATAAGCTGGTTTCTTATAGGAAGGCCATGTACCACACT 1204

Qy        150 LeuGlnValAlaSerSerArgAlaGlyLysLeuProValCysHisAspSerAspGlu 169
Db       1205 CTGGAGAAAGCCAGGGTTCGAGCTGGCAAGACCTTC----------TCC 1243

Qy        170 SerAspThrAlaLysAlaValGluValGlnAsnLysProMetIleGluTrpAlaLeuGly 189
```

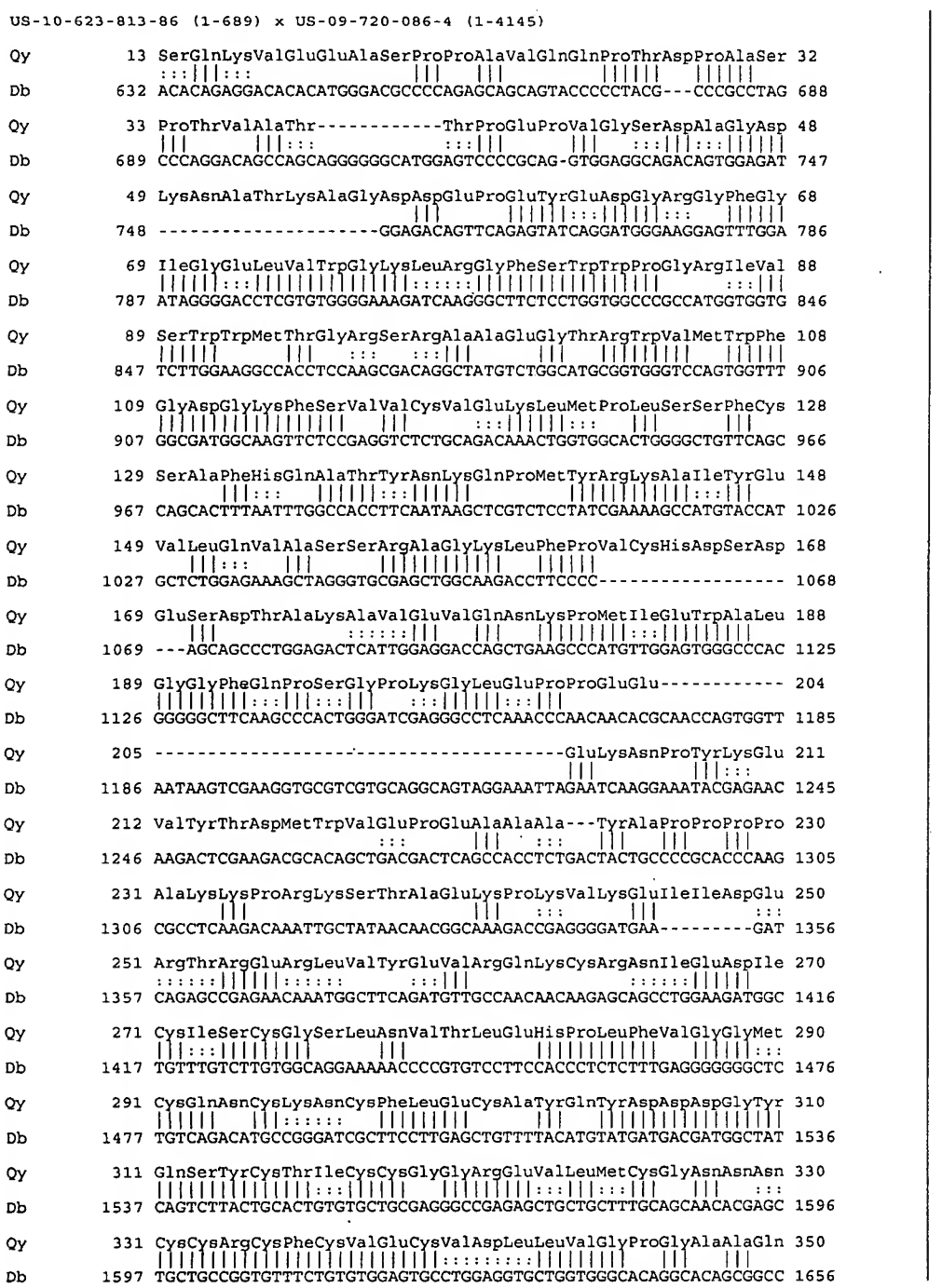
```
Db       1244 AGCAGTCCTGGAGAGTCACTGGAGGACCAGCTGAAGCCCATGCTGGAGTGGGCCCACGGT 1303

Qy        190 GlyPheGlnProSerGlyProLysGlyLeuGluPro---------------------- 201

Db       1304 GGCTTCAAGCCTACTGGGATCGAGGGCCTCAAACCCAACAAGAAGCAACCAGTGGTTAAT 1363

Qy        202 ----------------------------------------ProGluGluGluLysAsnPro 208

Db       1364 AAGTCGAAGGTGCGTCGTTCAGACAGTAGGAACTTAGAACCCAGGAGACGCGAGAACAAA 1423

Qy        209 TyrLysGluValTyrThrAspMetTrpValGluProGluAlaAlaAlaTyrAlaProPro 228

Db       1424 AGTCGAAGACGCACAACCAAT--------------GACTCTGCTGCTTCTGAGTCCCCC 1468

Qy        229 ProProAlaLysLysProArgLysSerThrAlaGluLysProLysValLysGluIleIle 248

Db       1469 CCACCCAAGCGCCTCAAGACAAATAGCTATGGCGGGAAGGACCGAGGGGAG--------- 1519

Qy        249 AspGluArgThrArgGluArgLeuValTyrGluValArgGlnLysCysArgAsnIleGlu 268

Db       1520 GATGAGGAGAGCCGAGAACGGATGGCTTCTGAAGTCACCAACAACAAGGGCAATCTGGAA 1579

Qy        269 AspIleCysIleSerCysGlySerLeuAsnValThrLeuGluHisProLeuPheValGly 288

Db       1580 GACCGCTGTTTGTCCTGTGGAAAGAAGAACCCTGTGTCCTTCCACCCCCTCTTTGAGGGT 1639

Qy        289 GlyMetCysGlnAsnCysLysAsnCysPheLeuGluCysAlaTyrGlnTyrAspAspAsp 308

Db       1640 GGGCTCTGTCAGAGTTGCCGGGATCGCTTCCTAGAGCTCTTCTACATGTATGATGAGGAC 1699

Qy        309 GlyTyrGlnSerTyrCysThrIleCysCysGlyGlyArgGluValLeuMetCysGlyAsn 328

Db       1700 GGCTATCAGTCCTACTGCACCGTGTGCTGTGAGGGCCGTGAACTGCTGCTGTGCAGTAAC 1759

Qy        329 AsnAsnCysCysArgCysPheCysValGluCysValAspLeuLeuValGlyProGlyAla 348

Db       1760 ACAAGCTGCTGCAGATGCTTCTGTGTGGAGTGTCTGGAGGTGCTGGTGGGCGCAGGCACA 1819

Qy        349 AlaGlnAlaAlaIleLysGluAspProTrpAsnCysTyrMetCysGlyHisLysGlyThr 368

Db       1820 GCTGAGGATGCCAAGCTGCAGGAACCCTGGAGCTGCTATATGTGCCTCCCTCAGCGCTGC 1879

Qy        369 TyrGlyLeuLeuArgArgArgGluAspTrpProSerArgLeuGlnMetPhePheAlaAsn 388

Db       1880 CATGGGGTCCTCCGACGCAGGAAAGATTGGAACATGCGCCTGCAAGACTTCTTCACTACT 1939

Qy        389 AsnHisAsp---GlnGluPheAspProProLysValTyrProProValProAlaGluLys 407

Db       1940 GATCCTGACCTGGAAGAATTTGAGCCACCCAAGTTGTACCCAGCAATTCCTGCAGCCAAA 1999

Qy        408 ArgLysProIleArgValLeuSerLeuPheAspGlyIleAlaThrGlyLeuLeuValLeu 427

Db       2000 AGGAGGCCCATTAGAGTCCTGTCTCTGTTTGATGGAATTGCAACGGGGTACTTGGTGCTC 2059

Qy        428 LysAspLeuGlyIleGlnValAspArgTyrIleAlaSerGluValCysGluAspSerIle 447

Db       2060 AAGGAGTTGGGTATTAAAGTGGAAAAGTACATTGCCTCCGAAGTCTGTGCAGAGTCCATC 2119

Qy        448 ThrValGlyMetValArgHisGlnGlyLysIleMetTyrValGlyAspValArgSerVal 467

Db       2120 GCTGTGGGAACTGTTAAGCATGAAGGCCAGATCAAATATGTCAATGACGTCCGGAAAATC 2179

Qy        468 ThrGlnLysHisIleGlnGluTrpGlyProPheAspLeuValIleGlyGlySerProCys 487

Db       2180 ACCAAGAAAAATATTGAAGAGTGGGGCCCGTTCGACTTGGTGATTGGTGGAAGCCCATGC 2239

Qy        488 AsnAspLeuSerIleValAsnProAlaArgLysGlyLeuTyrGluGlyThrGlyArgLeu 507

Db       2240 AATGATCTCTCTAACGTCAATCCTGCCCGCAAAGGTTTATATGAGGGCACAGGAAGGCTC 2299

Qy        508 PhePheGluPheTyrArgLeuLeuHisAspAlaArgProLysGluGlyAspAspArgPro 527

Db       2300 TTCTTCGAGTTTTACCACTTGCTGAATTATACCCGCCCCAAGGAGGGCGACAACCGTCCA 2359

Qy        528 PhePheTrpLeuPheGluAsnValValAlaMetGlyValSerAspLysArgAspIleSer 547

Db       2360 TTCTTCTGGATGTTCGAGAATGTTGTGGCCATGAAAGTGAATGACAAGAAAGACATCTCA 2419

Qy        548 ArgPheLeuGluSerAsnProValMetIleAspAlaLysGluValSerAlaAlaHisArg 567

Db       2420 AGATTCCTGGCATGTAACCCAGTGATGATCGATGCCATCAAGGTGTCTGCTGCTCACAGG 2479

Qy        568 AlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArgProLeuAlaSerThrValAsn 587

Db       2480 GCCCGGTACTTCTGGGGTAACCTACCCGGAATGAACAGGCCCGTGATGGCTTCAAAGAAT 2539

Qy        588 AspLysLeuGluLeuGlnGluCysLeuGluHisGlyArgIleAlaLysPheSerLysVal 607

Db       2540 GATAAGCTCGAGCTGCAGGACTGCCTGGAGTTCAGTAGGACAGCAAAGTTAAAGAAAGTG 2599

Qy        608 ArgThrIleThrThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisPheProVal 627

Db       2600 CAGACAATAACCACCAAGTCGAACTCCATCAGACAGGGCAAAAACCAGCTTTTCCCTGTA 2659

Qy        628 PheMetAsnGluLysGluAspIleLeuTrpCysThrGluMetGluArgValPheGlyPhe 647

Db       2660 GTCATGAATGGCAAGGACGACGTTTTGTGGTGCACTGAGCTCGAAAGGATCTTCGGCTTC 2719

Qy        648 ProValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuLeuGly 667

Db       2720 CCTGCTCACTACACGGACGTGTCCAACATGGGCCGCGGCGCCCGTCAGAAGCTGCTGGGC 2779

Qy        668 ArgSerTrpSerValProValIleArgHisLeuPheAlaProLeuLysGluTyrPheAla 687

Db       2780 AGGTCCTGGAGTGTACCGGTCATCAGACACCTGTTTGCCCCCTTGAAGGACTACTTTGCC 2839

Qy        688 Cys 688

Db       2840 TGT 2842
```

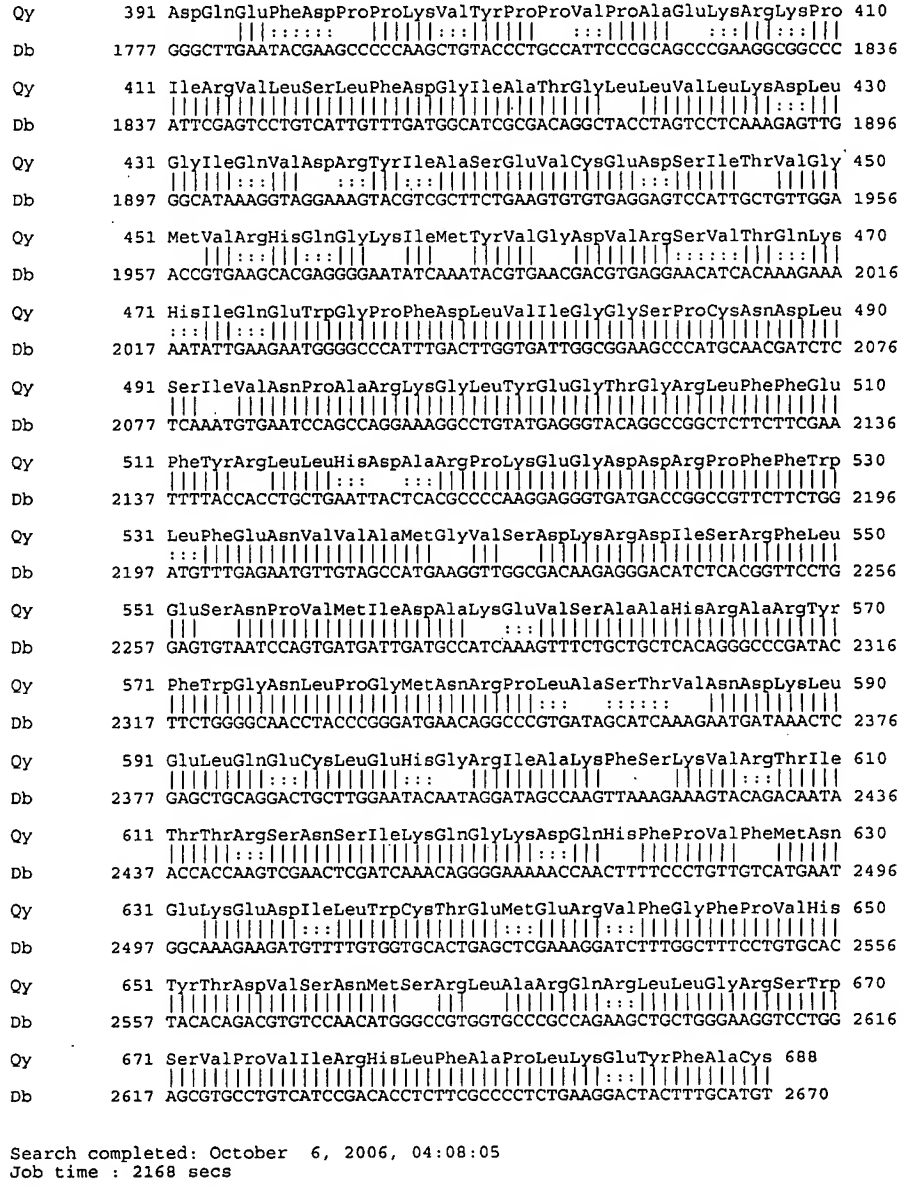
RESULT 15
US-09-720-086-4
; Sequence 4, Application US/09720086
; Publication No. US20060084053A1
; GENERAL INFORMATION:
; APPLICANT: Li, En
; APPLICANT: Okano, Masaki
; APPLICANT: Xie, Shaoping
; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,
; TITLE OF INVENTION: Polypeptides & Uses Thereof
; FILE REFERENCE: 0609.4560002
; CURRENT APPLICATION NUMBER: US/09/720,086
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/14373
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/090,906
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/093,993
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 4145
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-720-086-4

Alignment Scores:

Pred. No.:	1.23e-230	Length:	4145
Score:	2157.00	Matches:	409
Percent Similarity:	72.1%	Conservative:	94
Best Local Similarity:	58.6%	Mismatches:	155
Query Match:	57.3%	Indels:	41
DB:	5	Gaps:	7

US-10-623-813-86 (1-689) x US-09-720-086-4 (1-4145)





Search completed: October 6, 2006, 04:08:05
Job time : 2168 secs